



جامعة اليرموك

كلية التربية

قسم علم النفس الإرشادي والتربوي

أثر نسبة الفقرات ذات الأداء التفاضلي ومستواه في تقدير معالم فقرات الاختبار وقدرات
الأفراد وفق نموذج الاستجابة للفقرة ثلاثي المعلمة

**The Effect of Percentage of Items with Differential Functioning and its Level
on Estimating Item Parameters and Persons According to The Three
Parameters Model**

إعداد

علي صالح النوافلة

بإشراف

الأستاذ الدكتور أحمد يوسف قواسمه

حقل التخصص: القياس والتقويم

1434 هـ - 2013 م

أثر نسبة الفقرات ذات الأداء التفاضلي ومستواه في تقدير معالم فقرات الاختبار وقدرات الأفراد وفق نموذج الاستجابة للفقرة ثلاثي المعلمه

إعداد

علي صالح رجا نوافلة

بكالوريوس معلم صف، جامعة مؤتة، 2006.

ماجستير قياس وتقويم، جامعة اليرموك، 2008.

قدمت هذه الأطروحة استكمالاً لمتطلبات الحصول على درجة دكتوراه الفلسفة

في تخصص القياس والتقويم في جامعة اليرموك، إربد، الأردن.

وافق عليها

أحمد يوسف قواسمه رئيساً

أستاذ في القياس والتقويم، جامعة اليرموك

أحمد سليمان عوده عضواً

أستاذ في البحث والتقويم التربوي، جامعة جدارا

شفيق فلاح علاونة عضواً

أستاذ في علم النفس التربوي، جامعة اليرموك

ابراهيم فيصل رواشده عضواً

أستاذ في اساليب تدريس العلوم، جامعة اليرموك

نضال كمال الشريفين عضواً

أستاذ مشارك في القياس والتقويم، جامعة اليرموك

تاريخ مناقشة الأطروحة 2013/7/11

ب

الإهداء

إلى من أعطوا كل أنواع العطاء

إلى من بذلوا وجادوا بسخاء

إلى من أفنيا عمرهما في سبيل غيرهما دون ثناء

إلى من علماني أن أثبت حتى لو زال الجبل وانشق

والدي العزيزين

إلى التي جمعتني بها الأقدار فكانت خير رفيق ومعين ونبراس

ملاً دنياي نورا وسرورا

زوجتي الحبيبة

إلى من أفرح عند لقياهم، وبهم أستمد عزيمتي وإصراري

إلى أعلى الوجود

أبنائي "عامر وعبدالرحمن"

إلى من يشاركونني أفراحي، وكانوا للعين نورا وللقلب بهجاً وسروراً .

إخواني وأخواتي

إلى كل من علمني حرفاً وله الفضل علي.... أساتذتي.

إليهم جميعاً أهدي هذا الجهد العلمي المتواضع.

علي

شكر وتقدير

بسم الله الرحمن الرحيم

رَبِّ اَوْزَعِي اَنْ اَشْكُرَ نِعْمَتَكَ الَّتِي اَنْعَمْتَ عَلَیَّهَا وَ اَلَدِّيْ وَ اَنْ اَعْمَلَ صَالِحًا
تَرْضَاهُ وَ اَذْكُرْنِي بِرَحْمَتِكَ فِي عِبَادِكَ الصَّالِحِينَ

صدق الله العظيم سورة النمل آیه (19)

الحمد لله الذي منحني القدرة، وأمدني بالصبر، حتى تمكنت من إعداد هذه الأطروحة في صورتها الحالية، ولا يسعني بعد أن انتهيت من إعدادها إلا أن اعترف لكل ذي فضل عليّ بفضلته، فإن أهل الفضل والعطاء هم أهل للشكر والثناء .

أشكر أستاذي الفاضل الأستاذ الدكتور أحمد قواسمه، الذي منحني شرفاً عظيماً بالإشراف على هذه الأطروحة، وعاش معي متاعبها، فقد كان يوجهني ويعينني على البحث والتقصي، وبصرني بكثير مما خفي علي، ونور لي الطريق بغزارة علمه، وقدم لي النصح والإرشاد في كل خطوة من خطوات الأطروحة، فله مني كل الشكر والثناء والتقدير والاحترام.

كما أتقدم بالشكر الجزيل إلى أساتذتي الكرام، الأستاذ الدكتور أحمد عوده، والأستاذ الدكتور شفيق علاونة، والأستاذ الدكتور ابراهيم رواشده، والدكتور نضال الشريفين؛ الذين تكرموا بقبول مناقشة هذه الأطروحة، وتحملوا أعباء قراءتها، بابداء ملاحظاتهم القيمة وإغنائها بأرائهم ومقترحاتهم القيمة وصولاً بها إلى مصاف الرسائل العلمية، كما أقدم شكري وعظيم امتناني لأعضاء هيئة التدريس الذين تعلمت على أيديهم مواضيع وأسس القياس والتقويم التربوي.

فهرس المحتويات

الموضوع	رقم الصفحة
الإهداء.....	ج
الشكر والتقدير.....	د
فهرس المحتويات.....	هـ
قائمة الجداول.....	ز
قائمة الأشكال.....	ط
قائمة الملاحق.....	ي
الملخص باللغة العربية.....	ك
الفصل الأول: خلفية الدراسة ومشكلتها.....	1
مشكلة الدراسة وأسئلتها.....	6
أسئلة الدراسة.....	7
أهمية الدراسة.....	8
التعريفات الاصطلاحية والاجرائية.....	9
متغيرات الدراسة.....	10
محددات الدراسة.....	12
الفصل الثاني: الاطار النظري والدراسات السابقة.....	13

48.....	الفصل الثالث: الطريقة والإجراءات.....
48.....	ظروف توليد البيانات
48.....	عينة الدراسة.....
49.....	إجراءات توليد البيانات
56.....	المعالجات الإحصائية.....
57.....	الفصل الرابع: النتائج ومناقشتها.....
57.....	نتائج اجابة السؤال الأول.....
60.....	نتائج اجابة السؤال الثاني
64.....	نتائج اجابة السؤال الثالث
74.....	نتائج اجابة السؤال الرابع
81	نتائج اجابة السؤال الخامس.....
87.....	الاستنتاجات
88.....	التوصيات
89.....	قائمة المراجع
94.....	الملاحق
103.....	الملخص باللغة الإنجليزية

قائمة الجداول

رقم الجدول	المحتوى	الصفحة
1	قيم معالم الفقرات ثنائية التدرج من حيث (الصعوبة والتمييز والتخمين) لدى المجموعة المرجعية تبعاً للنموذج اللوجستي ثلاثي المعلمة.	50
2	أرقام الفقرات ذات الأداء التفاضلي وفقاً لنسبتها في الاختبار (10%,20%,40%).	51
3	نسبة الفقرات ذات الأداء التفاضلي، ومستوى الأداء التفاضلي للفقرة لدى المجموعات المستهدفة.	53
4	أعداد وأرقام الأفراد غير المطابقين في كل مجموعة وفقاً لنسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة، ونتائج الفرق بين نسب الأفراد المطابقين عبر جميع المجموعات.	58
5	إحصائيات مطابقة الفقرات للمجموعات التي ظهر عندها فقرات غير مطابقة وفقاً لنسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة.	59
6	نتائج التحليل العاملي لفقرات الإختبار النهائية في كل مجموعة وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي.	61
7	الإحصائيات الخاصة بمؤشر ZQ3 للإستقلال الموضعي في كل مجموعة من المجموعات المستهدفة والمجموعة المرجعية وفقاً لنسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي.	62
8	مؤشرات الإستقلال الموضعي في كل مجموعة من المجموعات المستهدفة والمجموعة المرجعية وفقاً لـ إلت ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة.	63
9	الوسط الحسابي والانحراف المعياري لمعلمة الصعوبة لفقرات الاختبار في كل مجموعة حسب نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.	65
10	نتائج تحليل التباين الثنائي للمتوسطات الحسابية لمعلمة الصعوبة حسب نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.	66
11	نتائج اختبار Games-Howell للكشف عن مواقع الفروق بين المتوسطات الحسابية لمعلمة الصعوبة وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار.	67

12	الوسط الحسابي والانحراف المعياري لمعلمة التمييز لفقرات الاختبار في كل مجموعة حسب نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.	68
13	نتائج تحليل التباين الثنائي للمتوسطات الحسابية لمعلمة التمييز لفقرات الاختبار حسب نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة.	69
14	نتائج اختبار Scheffe للكشف عن مواقع الفروق بين المتوسطات الحسابية لمعلمة التمييز لفقرات الاختبار وفقاً لمستوى الأداء التفاضلي للفقرة.	69
15	الوسط الحسابي والانحراف المعياري لمعلمة التخمين لفقرات الاختبار في كل مجموعة حسب نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.	71
16	نتائج تحليل التباين الثنائي للمتوسطات الحسابية لمعلمة التخمين لفقرات الاختبار حسب نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة.	72
17	نتائج اختبار Scheffe للكشف عن مواقع الفروق بين المتوسطات الحسابية لمعلمة التخمين لفقرات الاختبار وفقاً لمستوى الأداء التفاضلي للفقرة.	72
18	الأوساط الحسابية والانحرافات المعيارية لتقديرات قدرة الأفراد على الاختبار في كل مجموعة وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.	75
19	نتائج تحليل التباين الثنائي لتقديرات قدرات الأفراد على الاختبار وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.	76
20	الأوساط الحسابية والانحرافات المعيارية للعلامات الملاحظة للأفراد على الاختبار في كل مجموعة وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.	77
21	نتائج تحليل التباين الثنائي للعلامات الملاحظة للأفراد على الاختبار وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.	78
22	نتائج اختبار Games-Howell للكشف عن مواقع الفروق للعلامات الملاحظة للأفراد على الاختبار وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار.	79
23	نتائج اختبار Games-Howell للكشف عن مواقع الفروق للعلامات الملاحظة للأفراد على الاختبار وفقاً لمستوى الأداء التفاضلي للفقرة.	79
24	نتائج الإحصائي M للكشف عن جوهرية الفرق بين معاملي الثبات الأمبريقي للاختبار في كل مجموعة من المجموعات المستهدفة مقارنة مع المجموعة المرجعية.	82
25	الوسط الحسابي والانحراف المعياري لكمية المعلومات التي يقدمها الاختبار في كل مجموعة وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.	83

قائمة الأشكال

الصفحة	الشكل
16	الشكل:1 معالم الفقرة
27	الشكل:2 منحنى الفقرة الذي يظهر أداءً تفاضلياً منتظماً بين المجموعات.
28	الشكل:3 منحنى الفقرة الذي يظهر أداءً تفاضلياً غير منتظم بين المجموعات.
80	الشكل:4 التفاعل بين نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة وفقاً للعلامات المشاهدة للأفراد على الاختبار.
84	الشكل:5 رسوم منحنيات دالة معلومات الاختبار مع الخطأ المعياري له في كل مجموعة من المجموعات المستهدفة مقارنةً مع المجموعة المرجعية وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.

قائمة الملاحق

الصفحة

الملحق

- (أ):قيم معالم الفقرة (الصعوبة والتمييز والتخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المرجعية.....94
- (ب):قيم معالم الفقرة (الصعوبة والتمييز والتخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المستهدفة التي تشتمل على (10% من الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة 0.5).....95
- (ت):قيم معالم الفقرة (الصعوبة والتمييز والتخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المستهدفة التي تشتمل على (10% من الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة 1.5).....96
- (ث):قيم معالم الفقرة (الصعوبة والتمييز والتخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المستهدفة التي تشتمل على (20% من الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة 0.5).....97
- (ج):قيم معالم الفقرة (الصعوبة والتمييز والتخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المستهدفة التي تشتمل على (20% من الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة 1.5).....98
- (ح):قيم معالم الفقرة (الصعوبة والتمييز والتخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المستهدفة التي تشتمل على (40% من الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة 0.5).....99
- (خ):قيم معالم الفقرة (الصعوبة والتمييز والتخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المستهدفة التي تشتمل على (40% من الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة 1.5).....100
- (د):رسوم منحني خصائص الفقرة لكل فقرة من فقرات الاختبار في كل مجموعة من المجموعات المستهدفة، بالإضافة إلى المجموعة المرجعية وفقاً لمتغيري الدراسة.....101

المخلص

النوافلة، علي صالح. أثر نسبة الفقرات ذات الأداء التفاضلي ومستواه في تقدير معالم فقرات الاختبار وقدرات الأفراد وفق نموذج الاستجابة للفقرات ثلاثي المعلمة. إطروحة دكتوراة، جامعة اليرموك، 2013. (المشرف: أ.د. أحمد يوسف قواسمه).

هدفت هذه الدراسة إلى بيان أثر نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرات في تقدير معالم فقرات الاختبار وقدرات الأفراد وفق نموذج الاستجابة للفقرات ثلاثي المعلمة باستخدام بيانات مولدة، حيث تضمن تصميم الدراسة ظروفاً مختلفة بناءً على متغيري الدراسة، وهما (نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرات)، فقد كانت نسبة الفقرات ذات الأداء التفاضلي الموجودة في الاختبار هي (10%, 20%, 40%)، وكان مستوي الأداء التفاضلي للفقرات هما: (مستوى متدنٍ بمقدار 0.5 ومستوى عالٍ بمقدار 1.5). وبناءً على ذلك تكون تصميم الدراسة من (6) مجموعات مستهدفة تختلف فيما بينها بنسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرات، بالإضافة إلى المجموعة المرجعية التي لا تشتمل على أي فقرات ذات أداء تفاضلي، والتي تمثل القاعدة الأساسية للمقارنة، مما نتج (7) مجموعات للدراسة.

وقد تم استخدام برنامج (Wingen) لتوليد استجابات (510) أفراد في كل مجموعة، وقد تكونت العينة الكلية من (3570) مفحوصاً، ومن ثم تم استخدام برنامج (Bilog MG) للتحقق من افتراض أحادية البعد، ومطابقة كل من الفقرات والأفراد للنموذج، وتقدير معالم الفقرات والمفحوصين، وحساب دالة معلومات الاختبار والأخطاء المعيارية في تقدير القدرة، وذلك حسب كل مجموعة من المجموعات السابقة. وقد أشارت النتائج بوجود فروق ذات دلالة إحصائية بين متوسطات تقديرات معالم الصعوبة للفقرات تعزى إلى نسبة

الفقرات ذات الأداء التفاضلي، ولصالح المجموعة المستهدفة التي تشتمل على (40% من الفقرات ذات الأداء التفاضلي)، كما وتوجد فروق ذات دلالة إحصائية بين متوسطات تقديرات معالم التمييز ومعالم التخمين للفقرات تُعزى إلى مستوى الأداء التفاضلي للفقرة، ولصالح المجموعة المرجعية، وفيما يتعلق بقدرات الأفراد بيّنت النتائج عدم وجود فروق ذات دلالة إحصائية بين متوسطات تقديرات معالم قدرات الأفراد تُعزى إلى متغيري الدراسة، ويوجد فروق ذات دلالة إحصائية بين متوسطات العلامات الملاحظة للأفراد تُعزى إلى متغيري الدراسة (نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة) والتفاعل بينهما، وأما بالنسبة لكمية المعلومات التي يقدمها الاختبار فقد بيّنت النتائج وجود انخفاض في كمية المعلومات التي يقدمها الاختبار كلما زادت نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة.

الكلمات المفتاحية: نسبة الفقرات ذات الأداء التفاضلي في الاختبار، مستوى الأداء التفاضلي للفقرة، المجموعة المرجعية، المجموعة المستهدفة، معالم الفقرات، دالة معومات الاختبار، القدرة.

الفصل الأول

خلفية الدراسة ومشكلتها

المقدمة

لقد تركزت جهود الباحثين في بناء الاختبارات وتطويرها على استخراج خصائص الفقرات من حيث الصعوبة والتمييز والتخمين، وبالرغم من أهمية هذه المعالم إلا أنها ليست كافية للحكم على صلاحية فقرات الاختبار للغرض الذي أعدت لأجله، وذلك لأن الاستجابة لفقرات الاختبار قد تتأثر بعوامل أخرى بالإضافة إلى قدرة المفحوصين، كالجنس أو العرق أو مكان السكن أو اللغة أو المستوى الاقتصادي والاجتماعي بما يؤثر سلباً على نتيجة الاختبار، وبالتالي على القرارات المبنية عليه. وبناءً على ذلك توصف الفقرات بأنها ذات أداء تفاضلي نحو مجموعة أو فئة دون أخرى (Jensen, 1980).

وقد لاحظ تان وجيرل (Tan & Jeral, 2005) في عملية التحقق من صدق واستدلالات قدرات الأفراد في اختبار ما، أن أحد الخطوات الرئيسية في هذه العملية هو أن يتم تقييم درجة عدالة الاختبار من خلال تحليل الأداء التفاضلي للفقرة، حيث أن تحليل الأداء التفاضلي للفقرة هو إحدى الإجراءات الإحصائية المستخدمة لتحديد فيما إذا كانت فقرات الاختبار عادلة ومناسبة لتقييم المعارف لمبحث دراسي معين عبر مجموعات متشابهة من المفحوصين من نفس مستوى القدرة. وهذا التحليل مبني على الافتراض بأن الأفراد المتقدمين للاختبار يمتلكون معارف متشابهة، ويجب أن يكون أدائهم على فقرات الاختبار الفردية متشابهاً بغض النظر عن متغيرات الجنس أو العرق أو غيرها من الخصائص. ويتفق "تان وجيرل" مع الدراسات السابقة بأن تحليل الأداء التفاضلي للفقرة يجب إجراؤه من أجل الكشف عن الفروق بين المجموعات من خلال تقييم احتمالية أن يجيب المفحوصون الذين يمتلكون قدرة متساوية على الفقرة بشكل صحيح. وبهذا

فإن وجود الأداء التفاضلي للفقرة مؤشر بأن هنالك عوامل مرتبطة تنتمي إلى مجموعة معينة تؤثر في احتمالية الإجابة الصحيحة للفقرة، وبالتالي فإنها تهدد التقييم العادل بين الأفراد. ولذلك فقد تم تطوير الإجراءات والطرق الإحصائية من أجل الكشف عن الأداء التفاضلي للفقرة في محاولة لإيجاد المساواة بين المفحوصين الذين يمتلكون نفس القدرة على السمة الكامنة التي تهدف فقرة الاختبار لقياسها.

وذكر عوده (2004) بعض الآثار السلبية للاختبارات، منها: التحيز لعينة الفقرات أو لصياغة الفقرات وخصائصها لطالب دون آخر في استخدامها لمفردات غير شائعة أو غامضة، والتحيز في محتوى الفقرات، يكون لصالح مجتمعات معينة، وخاصة الاختبارات ذات المعايير التي يتم نقلها أو ترجمتها لاستخدامها في مجتمعات أخرى. كما أن احتواء الاختبار على فقرات متحيزة يعكس أثراً سلبية على خصائص الفقرات، وأهم هذه الخصائص القدرة التمييزية للفقرة، بمعنى صدق الفقرة، والتي يترتب عليها انتهاك لمبدأ العدالة والمساواة بين المجموعات التي يقع عليها الاختبار فيما يتعلق بتوفير التغذية الراجعة أو باتخاذ القرار، أو القبول في برامج مختلفة. وقد يُعزى التحيز في فقرات الاختبار إلى عدة مصادر يمكن أن تكون لصالح فئات في المجتمع الواحد، مثل التحيز للجنس، أو للموقع الجغرافي، أو العرقي، أو الطبقي (المستوى الاجتماعي-الاقتصادي).

وقد أعتبر تحيز الفقرة مشكلة في السياقات التعليمية والقانونية والمهنية، مما جعله الموضوع الرئيس في القياس النفسي والتربوي خلال الأربعين سنة الماضية، وبرز كقضية أساسية من أهم قضايا الاختبارات. ويُلاحظ المتتبع للجهود التي بذلها علماء القياس خلال العقدين الماضيين أن هنالك إهتماماً نظرياً وعملياً بدراسة خاصية التحيز، ولقد أشار بيرك (Berk, 1982) إلى العديد من الأسباب الكامنة وراء دراسة التحيز منها:

1. تطوير طرق بناء اختبارات خالية من التحيز ما أمكن، وذلك باستبعاد الفقرات المتحيزة للمجموعات الفرعية للمفحوصين.

2. مراجعة الاختبارات الموضوعية وتطويرها للتخلص من الفقرات المتحيزة.

3. تطوير الطرق المستخدمة للكشف عن التحيزة للاختبارات وللفقرات واستخدام طرق جديدة.

وقد أدى هذا الاهتمام إلى اعتبار أن خلو الاختبار أو فقراته من الأداء التفاضلي من أهم الشروط التي ينبغي توافرها في الاختبار قبل نشره، وقد وضعت الجمعية الأمريكية للبحث التربوي (American Educational Research Association AERA, 1999) الأداء التفاضلي ضمن قواعد النشر للاختبارات، وأشارت إلى أنه إذا استخدمت الاختبارات في اتخاذ القرارات، فلا بد من الحصول على أدلة تجريبية واضحة من أجل دعم الإدعاءات بأن فقرات الاختبار لا تعكس أداءً تفاضلياً للفقرات عبر المجموعات من المفحوصين. كما وضعت الجمعية الأمريكية لعلم النفس (American Psychological Association, APA, 1988) الأداء التفاضلي أيضاً ضمن قواعد النشر، ونصت على أنه يجب على مطوري الاختبارات أن يشيروا إلى طبيعة الأدلة المتعلقة بملاءمة الاختبار للمجموعات ذات الخلفيات العرقية أو اللغوية المختلفة، كما يجب دراسة أداء الخاضعين للاختبار من أعراق مختلفة وبأحجام عينات كافية للتأكد من أن الفروقات في الأداء ترتبط بشكل أساسي بالمهارات التي يجري تقييمها أكثر من كونها ترتبط بعوامل ليست ذات صلة بالاختبار.

ويذكر كامب وغروجر (Kamp & Gruijter, 2005) أنه إذا حدث أو لوحظ الأداء التفاضلي للفقرات فمن المهم عندئذٍ تحديد أسباب حدوثه. هل الفقرات تُعتبر أقل ألفة للمجموعة المستهدفة من المجموعة المرجعية، وأن كان الأمر كذلك، فهل الفرق في الألفة غير متعلق بالقدرة التي نحن مهتمين بها؟ فإذا كان

الفرق بين أداء المجموعتين يُعزى فعلاً إلى عوامل ليست ذات علاقة بالقدرة فتكون الفقرة متحيزة، وهي بالطبع أمر غير مرغوب، وتعتبر خاصية خلو فقرات الاختبار من الأداء التفاضلي هي إحدى الخصائص المرغوبة.

وللتمييز بين الأداء التفاضلي والتحيز وعدالة الاختبار، فإن عدالة الاختبار تتحقق عندما يكون أداء المفحوصين الذين يمتلكون نفس المستوى من القدرة متماثلاً على فقرات الاختبار بغض النظر عن ثقافتهم أو جنسهم أو عرقهم. وتعد فقرة الاختبار متحيزة إذا أدت وظيفة مختلفة لدى مجموعة جزئية من المفحوصين في المجموعة الكلية، وعندما تكون لدى الأفراد المتساويين في القدرة فرص غير متساوية في احتمالية الإجابة الصحيحة للفقرة. ويوجد التحيز فقط عندما يكون الفرق في أداء المجموعتين غير حقيقي. والأداء التفاضلي شرط ضروري لكنه غير كافٍ للتحيز. فقد تكون الفقرة ذات أداء تفاضلي لكنها غير متحيزة عندما يُعزى فرق الأداء عليها للفروق الحقيقية في قدرات الأفراد للمجموعات التي تستجيب عليها. وبمعنى آخر يكون الأداء التفاضلي تحيزاً عندما يُعزى الفرق في الأداء إلى عوامل لا صلة لها في السمة المقاسة.

وقد تزايد الاهتمام لدى المؤسسات التعليمية المختلفة في بناء فقرات تتمتع بخصائص سيكومترية ملائمة، ولا تتأثر بخصائص عينة المفحوصين (القدرة)، بحيث يمكن تطبيقها على عينات مختلفة منهم دون أن تتحيز لمجموعة ضد أخرى، أو تطبيق عينات مختلفة من الفقرات على نفس عينة المفحوصين دون أن تتأثر الفقرة بخصائص عينة المفحوصين (القدرة). بالإضافة إلى الاهتمام بخلو الاختبار من مفردات (مصطلحات) غامضة أو غير شائعة من حيث البنية أو الشكل لمجموعة ما من المجموعات ضد مجموعة أخرى (Popham, 1999).

ولقد أشار فراري وزيمر (Frery & Zimmer, 1984) إلى أن احتواء الاختبار على فقرات ذات أداء تفاضلي له آثار سلبية على القرار المتخذ في ضوء علامة الاختبار وعلى بعض السمات الإحصائية لهذه العلامات؛ حيث ينخفض الصدق والثبات نتيجة لوجود فقرات ذات أداء تفاضلي في الاختبار.

ذكر زهان وزمبو (Zhen & Zumbo, 2009) أنه قد تم دراسة الأداء التفاضلي للفقرة بشكل كبير في حقل القياس التربوي والنفسي، وقد ركزت العديد من الدراسات بشكل أساسي على تعريفات وعلى الطرق المستخدمة للكشف عنه. وهناك قبول واسع بأن وجود الفقرات ذات الأداء التفاضلي يمكن أن يقلل من مستوى صدق أي اختبار ومع ذلك، ليست هناك معلومات كثيرة حول أثر الفقرات ذات الأداء التفاضلي على القرارات الإحصائية عندما يستخدم المحلل علامة الاختبار الملاحظة في تحليل البيانات وأي اختبارات أخرى لفحص الفرضيات. وبغض النظر عن كيفية وجود الفقرات ذات الأداء التفاضلي في الاختبار، فإنه لا يزال من غير الواضح آلية تأثير هذه الفقرات على النتائج والقرارات الإحصائية. ويرى كاماتا وشو (kamata & Chu, 2003) أن الوضع الأمثل يتمثل بإلغاء الفقرات ذات الأداء التفاضلي من الاختبار لضمان عدالته، وهذه الطريقة أبسط وأقل إثارة للجدل من استخدام المعلومات التي توفرها الفقرات ذات الأداء التفاضلي.

ولقد أجريت العديد من الدراسات المحلية على الأداء التفاضلي للفقرات، وقد كان الأداء التفاضلي للفقرة محور تركيز العديد من الدراسات السابقة التي تم إجراؤها في حقل القياس التربوي النفسي. مع ذلك، فإن معظم الدراسات السابقة مثل: دراسة الحمادنة (2007)، وعبد العزيز (2006)، قد ركزت بشكل رئيسي على تعريف الأداء التفاضلي للفقرة، وطرق الكشف عن الأداء التفاضلي للفقرة في ضوء بعض المتغيرات: مثل مصادر الأداء التفاضلي (الجنس - العرق - مكان السكن - المستوى الاقتصادي والاجتماعي)، ونوع الأداء التفاضلي (منتظم - غير منتظم)، وحجم العينة وطول الاختبار. وهناك أيضاً بعض من الدراسات التي

تناولت أثر الفقرات ذات الأداء التفاضلي وحذفها على دقة معادلة الاختبار في ضوء بعض المتغيرات من حيث نوع المعادلة (رأسية- أفقية)، ومن حيث طرق المعادلة وحجم العينة وطول الاختبار مثل: دراسة ارحيل(2011) والحياصات(2011). وبشكل خاص وفي حدود علم الباحث لا توجد أي دراسة في الأدب السابق بحثت في أثر الأداء التفاضلي للفقرة على تقدير معالم فقرات الاختبار وقدرات الافراد.

مشكلة الدراسة وأسئلتها

هناك الكثير من الأبحاث التي أجريت للكشف عن الأداء التفاضلي للفقرة، فبعضها يعتمد على النظرية الكلاسيكية وبعضها يعتمد على النظرية الحديثة في القياس. ومعظم هذه الأبحاث ركزت على المقارنة بين الطرق المختلفة للكشف عن الأداء التفاضلي للفقرة. وهناك قبول واسع بأن وجود الأداء التفاضلي في الفقرة يؤثر على تقدير معلمة الصعوبة على مستوى الفقرة لأن الفقرة التي تظهر أداءً تفاضلياً تكون أصعب لمجموعة من مجموعة أخرى، أما بالنسبة لوجود عدة فقرات ذات الأداء التفاضلي في الاختبار فإنه من غير الواضح حتى الآن آلية تأثير هذه الفقرات ذات الأداء التفاضلي على تقدير معالم الصعوبة للفقرات على مستوى الاختبار. بالإضافة لذلك، ليست هناك معلومات حول أثر الفقرات ذات الأداء التفاضلي الموجودة في الاختبار على تقديرات المعالم الأخرى لفقرات الاختبار وهي معلمة التمييز والتخمين، وبغض النظر عن كيفية وجود الفقرات ذات الأداء التفاضلي في الاختبار، فإنه لا يزال من غير الواضح آلية تأثير هذه الفقرات على معالم الفقرات في الاختبار (الصعوبة والتمييز والتخمين) وعلى معالم القدرات لدى الأفراد.

ونظراً للإهتمام المتزايد بنظرية الاستجابة للفقرة (Item Response Theory, IRT)، فإن هذه

الدراسة جاءت لتبين أثر نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة

على تقديرات معالم الفقرات للاختبار (الصعوبة والتمييز والتخمين)، ومعالم القدرة للأفراد، ومطابقة الفقرات والأفراد للنموذج، وكمية معلومات الاختبار .

ومن خلال التعريف بمشكلة هذه الدراسة، يمكن صياغة بعض الاسئلة التي هدفت هذه الدراسة الإجابة عنها على النحو الآتي:

1. ما مدى مطابقة الأفراد ومطابقة الفقرات للنموذج اللوجستي ثلاثي المعلمة تبعاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة؟

2. ما مدى تحقق افتراضات النموذج (احادية البعد والاستقلال الموضوعي) في الاختبار تبعاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة؟

3. هل توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين متوسطات تقديرات معالم فقرات الاختبار (الصعوبة والتمييز والتخمين) تُعزى لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة أو التفاعل بينهما؟

4. هل توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين متوسطات تقديرات قدرات الأفراد على الاختبار تُعزى لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة أو التفاعل بينهما؟

5. هل يوجد اختلاف في كمية معلومات الاختبار يُعزى لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة ؟

أهمية الدراسة

تعمل هذه الدراسة على محدد من أهم محددات بناء الاختبارات، حيث أن موضوع الأداء التفاضلي للفقرة يهدد صدق البناء والثبات وتقدير المعالم للقدرة والفقرة، وبالتالي تلوث النتائج، وبذلك يصعب اتخاذ قرارات تربوية في ضوء بيانات غير صادقة.

وتبرز أهمية الدراسة بسبب قلة الدراسات التي تناولت تأثير الفقرات ذات الأداء التفاضلي الموجودة في الاختبار ومستوى الأداء التفاضلي للفقرة على تقديرات معالم الفقرات وقدرات الأفراد وفقاً لنموذج الاستجابة للفقرة ثلاثي المعلمة. حيث أن أكثر الدراسات السابقة اعتمدت على احصائيات النظرية الكلاسيكية، أما الدراسة الحالية فهي تعتمد على نظرية الاستجابة للفقرة.

ومن الناحية العملية ، فإن الدراسة تستخدم بيانات مولدة، وهذا يضمن مطابقة البيانات للنموذج المستخدم، ويضمن أيضاً كون مؤشرات معالم الفقرات والقدرة حقيقية، وليست مقدرة، وكل ما سبق يعطي أهمية للنتائج بحيث تساعد المؤسسات والجهات التي تقوم ببناء وتطوير الاختبارات والمهتمين بموضوع الأداء التفاضلي للفقرة في وضع آلية لتقييم نتائجهم عند اشتغال اختباراتهم على فقرات ذات أداء تفاضلي على اختلاف نسبتها في الاختبار ومستواها في الحكم على مدى حجم التأثير للفقرات ذات الأداء التفاضلي في اختباراتهم.

وتعتبر الدراسة تطبيقاً عملياً للإثبات النظري لمقدار حجم الأثر الذي ينتج من وجود الفقرات ذات الأداء التفاضلي على اختلاف نسبتها في الاختبار على احصائيات نظرية الاستجابة للفقرة وهي: مطابقة الفقرات والافراد للنموذج المستخدم، وعلى افتراضات النموذج (احادية البعد والاستقلال الموضعي)، وتقديرات معالم الفقرات للاختبار (الصعوبة والتمييز والتخمين)، وتقديرات قدرات الأفراد وعلاماتهم الملاحظة، وكمية المعلومات التي يقدمها الاختبار.

التعريفات الاصطلاحية والاجرائية

نسبة الفقرات ذات الأداء التفاضلي: هي عدد الفقرات ذات الأداء التفاضلي الموجودة في الاختبار مقسوماً على عدد الفقرات الكلية في الاختبار، حيث كان هنالك ثلاث حالات للاختبار حسب أعداد الفقرات ذات الأداء التفاضلي في الاختبار، وهي (5,10,20) فقرة من أصل العدد الكلي لفقرات الاختبار والبالغ عدد فقراته (50) فقرة، وكانت هذه الأعداد تشكل النسب التالية للفقرات ذات الأداء التفاضلي الموجودة في الاختبار، وهي (10%,20%,40%).

مستوى الأداء التفاضلي للفقرة: هو مقدار الزيادة في قيمة معلمة الصعوبة للفقرة في المجموعات المستهدفة، حيث كان في هذه الدراسة مستويان لمقدار الزيادة في قيمة معلمة الصعوبة هما: مستوى متدنٍ بزيادة مقدارها (0.5) في قيمة معلمة الصعوبة لدى المجموعة المستهدفة، ومستوى عالٍ بزيادة مقدارها (1.5) في قيمة معلمة الصعوبة لدى المجموعة المستهدفة.

المجموعة المرجعية: هي المجموعة التي لا تشتمل على أي فقرة ذات أداء تفاضلي، وهذه المجموعة كانت بمثابة القاعدة الأساسية للمقارنة مع باقي المجموعات المستهدفة.

المجموعة المستهدفة: هي المجموعة التي تشتمل على فقرات ذات أداء تفاضلي، ونظراً لوجود ثلاث حالات للاختبار حسب نسب الفقرات ذات الأداء التفاضلي وهي (10%,20%,40%)، ومستويان للأداء التفاضلي للفقرة (مستوى متدنٍ بمقدار 0.5، ومستوى عالٍ بمقدار 1.5)؛ فقد كان هنالك (6) مجموعات مستهدفة تختلف فيما بينها بنسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة.

صعوبة الفقرة: هي النقطة على متصل القدرة التي تقابل احتمال الإجابة الصحيحة $(1+C)/2$ في حال وجود التخمين نظراً لاستخدام النموذج اللوجستي ثلاثي المعلمة، وتأخذ صعوبة الفقرة قيمة من $(-\infty, +\infty)$ ، وعملياً من $(-3, +3)$.

تمييز الفقرة: هي قدره الفقرة على التمييز بين الطلبة من ذوي التحصيل المرتفع والتحصيل المنخفض، ويستدل عليها بمؤشر ميل المماس لمنحنى خصائص الفقرة عند النقطة التي تحدد صعوبتها.

التخمين: هو احتمالية إجابة الفقرة إجابة صحيحة عن طريق التخمين للأفراد من ذوي القدرات المنخفضة، ويمثل الخط التقاربي الأدنى للمنحنى.

دالة المعلومات الاختبار: هو إقتران ناتج عن جمع كمية المعلومات التي تُقدمها الفقرات عند مستوى قدرة محدد، حيث يُقدم منحنى معلومات الاختبار معلومات عن دقة الاختبار في تقدير قدرات الأفراد.

متغيرات الدراسة

المتغيرات المستقلة:

1. نسبة الفقرات ذات الأداء التفاضلي في الاختبار، ولها ثلاثة ظروف هي: (10%, 20%, 40%) من فقرات الاختبار ذات أداء تفاضلي.

2. مستوى الأداء التفاضلي للفقرة: كان في هذه الدراسة مستويان للأداء التفاضلي للفقرة، وهما: (مستوى متدنٍ بمقدار 0.5، ومستوى عالٍ بمقدار 1.5).

وبناءً على ذلك؛ فقد تكون (7) مجموعات مستقلة للدراسة، وهي:

1. المجموعة المرجعية، والتي لا تشتمل على أي فقرة ذات أداء تفاضلي، حيث تمثل هذه المجموعة القاعدة الأساسية للمقارنة مع باقي المجموعات.
2. المجموعة المستهدفة التي تشتمل على (10%) من فقرات الاختبار ذات أداء تفاضلي، ومستوى الأداء التفاضلي (0.5).
3. المجموعة المستهدفة التي تشتمل على (10%) من فقرات الاختبار ذات أداء تفاضلي، ومستوى الأداء التفاضلي (1.5).
4. المجموعة المستهدفة التي تشتمل على (20%) من فقرات الاختبار ذات أداء تفاضلي، ومستوى الأداء التفاضلي (0.5).
5. المجموعة المستهدفة التي تشتمل على (20%) من فقرات الاختبار ذات أداء تفاضلي، ومستوى الأداء التفاضلي (1.5).
6. المجموعة المستهدفة التي تشتمل على (40%) من فقرات الاختبار ذات أداء تفاضلي، ومستوى الأداء التفاضلي (0.5).
7. المجموعة المستهدفة التي تشتمل على (40%) من فقرات الاختبار ذات أداء تفاضلي، ومستوى الأداء التفاضلي (1.5).

المتغيرات التابعة:

1. مدى مطابقة الفقرات ومطابقة الأفراد للنموذج ثلاثي المعلمة.

2. افتراضات النموذج (احادية البعد والاستقلال الموضوعي).

3. معالم الفقرات وهي: (الصعوبة والتمييز والتخمين).

4. معالم القدرة للأفراد وعلاماتهم الملاحظة.

5. كمية معلومات الاختبار.

محددات الدراسة:

1. اقتصرت الدراسة على النموذج اللوجستي ثلاثي المعلمة؛ لأنه يضم جميع المعالم الممكنة للفقرة .

2. إقتصرت الدراسة على ثلاث ظروف لنسب الفقرات ذات الأداء التفاضلي في الاختبار، وهي (10%,20%,40%).

3. اقتصرت الدراسة على مستويين للأداء التفاضلي للفقرة، وهما (مستوى متدنٍ بمقدار 0.5 ، ومستوى عالٍ بمقدار 1.5).

4. اقتصرت الدراسة على فقرات من نوع الأداء التفاضلي المنتظم (UNIFORM).

5. اقتصرت الدراسة على فقرات ثنائية التدرج (0,1).

6. اقتصرت الدراسة على استخدام بيانات مولدة (Simulated Data).

الفصل الثاني

الأطار النظري والدراسات السابقة

الإطار النظري للدراسة.

يتناول الأطار النظري لهذه الدراسة توضيح لكل من المفاهيم الأساسية ذات الصلة بموضوع الدراسة، والتي تتمركز حول نظرية الاستجابة للفقرة، والأداء التفاضلي للفقرة بعد أن تم تمييزه عن التحيز كما ورد في المقدمة.

أولاً : نظرية الاستجابة للفقرة: (IRT) Item Response Theory.

لقد استخدمت نظرية الاختبار الكلاسيكية (CTT) Classical Test Theory بشكل واسع لتفسير التفاعلات بين استجابات الفرد وبين فقرات الاختبار. واعتبرت هذه النظرية أن العلامة الملاحظة مركبة من علامة حقيقية وعلامة خطأ (Lord & Nonvick, 1968). ورغم شعبيتها فقد عانت نظرية الاختبار العديد من المشكلات مما ساهم في ظهور نظرية الاستجابة للفقرة. فقد كانت جوانب القصور الأول الذي عانت منه هذه النظرية هو أن صعوبة الفقرة وتمييز الفقرة أعمدتا على عينة المفحوصين. والجانب الثاني اعتماد نظرية الاختبار في مقارنتها للمفحوصين على القدرة المقاسة بنفس فقرات الاختبار أو أشكال متوازية للاختبار. وجانب القصور الثالث لنظرية الاختبار تتمثل في ثبات الاختبار الذي عرّف بالصور المتوازية والذي يتطلب تطبيق ثاني لنفس الاختبار، وأخيراً افتراضاتها بأن هنالك خطأ قياس واحد عند جميع مستويات القدرة (Hambleton & Swaminathan, 1985)، وبالرغم من أن كل من النظريتين الكلاسيكية والحديثة قد سعيا لحل المشاكل التي تظهر في القياس النفسي والتربوي إلا أن النظرية الحديثة تميزت عن النظرية

الكلاسيكية لقيامها على افتراضات قوية، وتمتلك نماذج تصف العلاقة بين قدرة المفحوصين واستجاباتهم اعتماداً على خصائص فقرات الاختبار.

وتفترض نظرية الاستجابة للفقرة أنه يمكن التنبؤ بأداء الأفراد في اختبار نفسي أو تربوي في ضوء سمة أو خاصية مميزة لهذا الأداء تدعى السمات الكامنة. ونظراً لصعوبة ملاحظة هذه السمات بطريقة مباشرة أو قياسها بصورة مباشرة فإنه يجب تقديرها أو الاستدلال عليها، من خلال استجابات الأفراد الملاحظة على فقرات الاختبار الذي يقيس تلك السمة أو القدرة المطلوبة، ولهذا أطلق عليها نظرية السمات الكامنة. وقد انبثق عن هذه النظرية مجموعة من النماذج تُعرف باسم نماذج السمات الكامنة (Latent Trait Models)، ولكل نموذج معادلة رياضية تحدد فيها العلاقة بين أداء الفرد على الفقرة والقدرة التي تكمن وراء هذا الأداء وتفسره. ونظراً للاستخدام المتزايد لنظرية الاستجابة للفقرة؛ فإن هنالك بعض المزايا التي أشار إليها هاملتون وسوا مينثان (Hambleton & Swaminathan, 1985) وهي :

1. إن تقديرات الخصائص السيكمترية للفقرات (الصعوبة والتمييز) مستقلة عن عينة الأفراد التي استخدمت في تقدير هذه الخصائص (Item Free).

2. إن تقديرات قدرات الأفراد مستقلة عن عينة الفقرات التي تطبق عليهم (Person Free).

3. توفر النظرية مؤشراً إحصائياً لتقدير درجة الدقة في قياس القدرة لكل فرد عن طريق دالة المعلومات .

وتستند نماذج نظرية الاستجابة للفقرة إلى بعض الافتراضات، والتي يجب تحقيقها في البيانات لكي تؤدي

إلى نماذج يمكن الوثوق بها، وهذه الافتراضات هي:

أولاً - أحادية البعد (Unidimensionality): والتي تعني أن أداء المفحوص على الاختبار يمكن أن

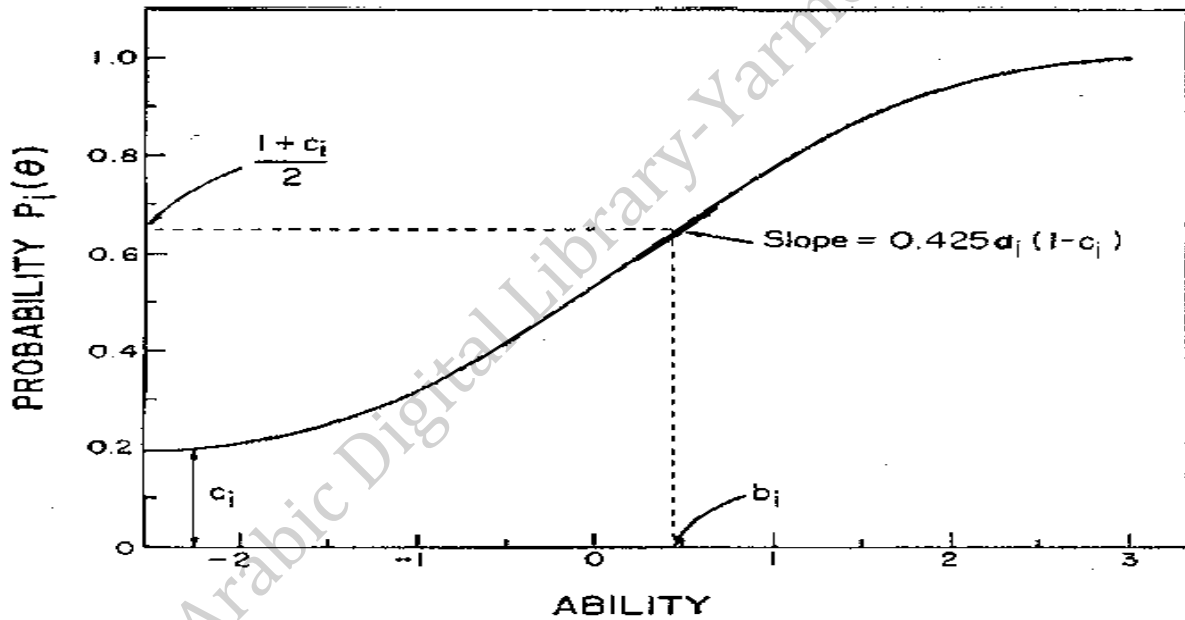
يعزى إلى سمة واحدة. وهذا الافتراض يصعب تحقيقه في كثير من الأحيان؛ لأن هناك عوامل أخرى ربما

تؤثر في الأداء إضافة إلى قدرة الفرد، كالعوامل الشخصية أو المعرفية أو ظروف تطبيق الاختبار. وتسمى الاختبارات التي تقوم على قياس سمة واحدة للمفحوصين اختبارات أحادية البعد.

ثانياً - **الاستقلال الموضعي لل فقرات (local Independence):** ويقصد بذلك أن استجابات المفحوص لفقرات الاختبار مستقلة إحصائياً عند مستوى محدد من القدرة، والتي تعني أن استجابة المفحوص على أي فقرة لا تُؤثر في استجابته على أي فقرة أخرى بالاختبار. ويقال بأن علامات الفقرات في هذه الحالة مستقلة إحصائياً. والاستقلال الموضعي لا يعني أن الفقرات لا ترتبط بعضها ببعض، بل على العكس يجب أن ترتبط الفقرات بارتباطات عالية، وذلك لتحقيق افتراض أحادية البعد في الاختبار. فإذا لم تكن الفقرات مستقلة إحصائياً أي غير مترابطة عند مستوى معين من القدرة، فهذا يعني بأن احتمالية الاستجابة الصحيحة لبعض المفحوصين على الاختبار ستكون أعلى مقارنة مع بعض أقرانهم عند نفس المستوى من القدرة. وبالتالي لابد أن تكون هنالك أكثر من قدرة لتفسير أداء المفحوص على الاختبار. وهذا إنتهاك لافتراض أحادية البعد وهذا لا يعني أن الفقرات غير مترابطة عبر المجموعة الكلية للمفحوصين، أي عند مستويات مختلفة من القدرة أو على متصل القدرة. ويوجد هنالك العديد من المؤشرات الإحصائية للتحقق من افتراض الاستقلال الموضعي وأحادية البعد.

ثالثاً - **التحرر من السرعة في الأداء (Speediness):** أي أن عامل السرعة لا يلعب دوراً في الإجابة عن فقرات الاختبار، بمعنى أن إخفاق الأفراد في الإجابة عن فقرات الاختبار يرجع إلى انخفاض قدرتهم، وليس إلى عامل السرعة.

رابعاً- منحنى خصائص الفقرة (Item Characteristic Curve : ICC): إن منحنى خصائص الفقرة هو تمثيل بياني لدالة رياضية تصف العلاقة بين احتمال الاستجابة الصحيحة للفقرة والقدرة التي يقيسها الاختبار المكون من مجموعة من الفقرات. ويرمز لهذه الدالة عادة بالرمز $P(\theta)$ أي احتمال إجابة المفحوص الذي قدرته (θ) على الفقرة (i) ، ويشير هذا المنحنى بأن احتمال إجابة الفقرة إجابة صحيحة يزداد بإزدياد القدرة للمفحوصين، والشكل التالي يوضح طبيعة هذا المنحنى:



الشكل (1): معالم الفقرة

وتحدد طبيعة شكل هذا المنحنى من خلال ثلاثة معالم هي:

1. صعوبة الفقرة (b_i) : وهي النقطة على متصل القدرة التي تقابل احتمال الإجابة الصحيحة على الفقرة (0.5) في حال عدم وجود التخمين، أو $(1+C)/2$ في حال وجود التخمين، وتأخذ صعوبة الفقرة قيمةً من $(-\infty, +\infty)$ وعمليةً من $(-3, +3)$.

2. تمييز الفقرة (a_i): ويشير إلى القوة التمييزية للفقرة (i)، ويعرف التمييز رياضياً بميل منحنى خصائص

الفقرة (ICC)، وأعلى تمييز للفقرة عندما تكون ($b=0$).

3. درجة التخمين (c_i): وتمثل احتمالية اجابة الفقرة اجابة صحيحة عن طريق التخمين للأفراد من ذوي

القدرات المتدنية، ويمثل الخط التقاربي الأدنى للمنحنى. وبناءً على هذه المعالم يوجد ثلاثة نماذج

لمنحنى خصائص الفقرة.

نماذج نظرية الاستجابة للفقرة: IRT Models

نظراً لتوفر عدة أشكال رياضية لمنحنيات خصائص الفقرة، ونتيجة لاستخدام هذه النظرية والعمل بها،

فقد تم تطوير نماذج مختلفة تصف العلاقة بين أداء المفحوصين على الاختبار والسمات الكامنة وراء هذا

الأداء باقتران رياضي لوغاريتمي. ومن هذه النماذج: النماذج اللوجستية، والنماذج التي تستخدم المنحنى

الطبيعي التراكمي، وتختلف النماذج اللوجستية باختلاف عدد المعالم في كل نموذج، وأشهر هذه النماذج

هي: (Hambleton & Swaminathan, 1985).

1. النموذج اللوجستي ثلاثي المعلم: (Three-Parameter Logistic Model).

في النموذج اللوجستي ثلاثي المعلمة تختلف المعالم الثلاث للفقرة (الصعوبة والتمييز والتخمين) ويُسمى

هذا النموذج بنموذج لورد. ويُعد هذا النموذج امتداداً للنموذج اللوغاريتمي ثنائي المعلم، حيث أضاف لورد

معلماً ثالثاً وأطلق عليه معلم الخط التقاربي الأدنى (Lower Asymptotic Line) أو معلم التخمين

(Guessing Parameter). وهذا المعلم يحدد احتمال أن يجيب فرد يفترض أن مستوى قدرته منخفض

لخفاضاً لا نهائياً، ومع هذا يجيب إجابة صحيحة عن غالبية فقرات الاختبار عن طريق التخمين.

ويُوصف بالمعادلة الرياضية التالية:

$$P_i(\theta) = c_i + (1 - c_i) \frac{1}{1 + e^{-D a_i(\theta - b_i)}} \dots\dots\dots (1)$$

حيث:

$P_i(\theta)$: احتمال أن يجيب المفحوص ذي القدرة (θ) على الفقرة (i) إجابة صحيحة.

D: ثابت التدرج ويساوي تقريباً 1.7.

b_i : معلم الصعوبة للفقرة (i)، وهي النقطة التي تقع على متصل القدرة التي تقابل احتمال الإجابة الصحيحة $(1+C)/2$.

a_i : معلم تمييز الفقرة، وهو ميل منحنى خصائص الفقرة عند إنعطاف المنحنى.

C_i : (معلم التخمين للفقرة) وهو الخط التقاربي السفلي لمنحنى خصائص الفقرة الذي يمثل احتمالية إجابة المفحوص من ذوي القدرة المتدنية إجابة صحيحة على الفقرة عن طريق التخمين.

2. النموذج اللوجستي ثنائي المعلم: (Two- Parameter Logistic Model).

تقترب معلمة التخمين للفقرة من الصفر في هذا النموذج ، وتختلف فيه معالم الصعوبة والتمييز، وهو ما يُلاحظ عادةً في بناء الاختبارات، إذ من الصعب إيجاد مجموعة من الفقرات تميز بدرجة واحدة بين مستويات السمة أو القدرة التي يقيسها اختبار معين. كما يفترض عدم تأثر إجابات المفحوصين بعامل التخمين،

ويسمى هذا النموذج بنموذج "بيرنيوم"، ويوصف بالمعادلة التالية:

$$P_i(\theta) = \frac{1}{1 + e^{-D a_i(\theta - b_i)}} \dots\dots\dots (2)$$

3. النموذج اللوجستي أحادي المعلمة: (One – Parameter Logistic Model).

يفترض هذا النموذج تساوي معلمة التمييز لجميع الفقرات، وتقترَب فيه معلمة التخمين للفقرات من الصفر وتختلف فيه معالم الصعوبة للفقرات. ويُعرف هذا النموذج بنموذج راش، وهو من أكثر النماذج شهرة واستخداماً، وذلك لسهولة تطبيقه، ويوصف بالمعادلة الرياضية التالية:

$$P_i(\theta) = \frac{1}{1 + e^{-D(\theta_i - b_i)}} \dots\dots\dots (3)$$

وتهدف نظرية الاستجابة للفقرة والنماذج الرياضية المرتبطة بها إلى تقدير جميع معالم الفقرة والقدرة، وكلما كان هنالك مواءمة (Fit) بين النموذج المستخدم ومجموعة البيانات أدى ذلك إلى الحصول على تقديرات دقيقة لهذه المعالم. ويتم ذلك من خلال عمليات التقدير حيثما وجدت علاقة تقارب ممكنة بين الاحتمالات المتوقعة للممتحنين والاحتمالات الواقعية الفعلية لأدائهم في كل مستوى من مستويات القدرة، مع الأخذ بعين الاعتبار أن تقديرات معالم الفقرة وتقديرات القدرة للممتحنين يجب أن تُعدل وتُراجع بصفة مستمرة حتى يمكن الحصول على أقصى اتساق ممكن بين التنبؤات على أساس تقديرات القدرة ومعالم الفقرة وبين البيانات الواقعية الفعلية للاختبار، ويتم ذلك من خلال برامج الحاسب الآلي المُعد لذلك (Hambleton & Swaminathan, 1985).

تقديرات المعالم في نظرية الاستجابة للفقرة .

في ظل نظرية الاستجابة للفقرة يعتمد احتمال الإجابة الصحيحة على الفقرة على معلمة القدرة للمفحوص ومعالم الفقرة: الصعوبة والتمييز والتخمين والتي تُعد مجهولة، والتي يمكن تقديرها من خلال استجابات المفحوصين على الاختبار. والهدف الأساسي في ضوء نظرية الاستجابة للفقرة عند تقديم الاختبار

للمفحوصين هو تحديد موقع المفحوص على مقياس القدرة، وإذا أمكن الحصول على قدرة كل مفحوص على مقياس مشترك للقدرة فبهذا يكون قد تحقق هدفان الأول: يمكن تقييم المفحوص في حدود القدرة التي يملكها، والثاني: مقارنة بين المفحوصين بأهداف إعطائهم علامات واتخاذ قرارات مختلفة في ضوء ذلك. وهناك عدة طرق تُستخدم لتقدير قدرة المفحوصين والمعالج اللوجستية للفقرات ذكرها هاملتون وسواميناثان (Hambleton & Swaminathan, 1985) منها :

أولاً: طرق تقدير الأرجحية العظمى (Maximum Likelihood Estimation Methods):

وتعتبر من أكثر طرق التقدير شهرة، وفي هذه الطريقة يتم تقدير المعالج من خلال إجراءات تعظيم الاحتمالية للمعلمة المراد تقديرها، عندما تتوفر معلومات عن العينة. وتعتمد هذه الطريقة على نمط إجابة المفحوص لمجموعة من الفقرات حيث تأخذ الإجابة إما واحداً أو صفراً. ومن محددات هذه الطريقة أنه عندما يُجيب الفرد على جميع الفقرات فإن قيمة (θ) المقدرة تساوي $(+\infty)$ ، وعندما يُجيب إجابة خاطئة على جميع الفقرات فإن قيمة (θ) المقدرة تساوي $(-\infty)$. ومن المحددات أيضاً أنه لا يمكن تقدير المعالج للفقرات عندما يجيب عليها جميع المفحوصين إجابة صحيحة أو لا يجيب عليها احد من المفحوصين إجابة صحيحة. ومن أجل التغلب على هذه المحددات فإنه يتم حذف الأفراد الذين يجيبون إجابات صحيحة أو إجابات خاطئة على جميع الفقرات، وحذف الفقرات التي أجاب الجميع عليها والتي لم يجيب الجميع عليها، وكذلك يمكن استخدام طريقة "بييز" بدلاً منها لتجنب هذه المحددات. ويوجد مجموعة من الطرق المستخدمة في التقدير أهمها:

1. طريقة تقدير الأرجحية العظمى المشتركة (Joint Maximum Likelihood Estimation)

:(Method)

ويمكن تطبيق هذه الطريقة في النماذج اللوجستية الأحادية والثنائية والثلاثية، ويتم من خلال هذه الطريقة تقدير معالم القدرة والفقرة معاً، حيث يتم في هذه الطريقة عملية تقدير المعالم بخطوتين، في الخطوة الأولى يتم اختيار قيم مبدئية لمعالم الأفراد، ويُفضل أن تحسب هذه القيم المبدئية بقسمة لوغاريتم النسبة بين عدد الإجابات الصحيحة وعدد الإجابات الخاطئة لكل مفحوص وتحول إلى قيم معيارية، ويتم التعامل مع هذه المعالم للقدرات على أنها معلومة. وفي الخطوة الثانية يتم التعامل مع معالم القدرة أنها معلومة، ويجري تقدير معالم الفقرات، وتكرر الخطوتان السابقتان حتى تكون القيم التقديرية لمعالم الفقرات والأفراد لا تتغير بعد الإنتهاء من الخطوتين في مرتين متتاليتين.

2. طريقة الأرجحية العظمى المشروطة (Conditional Maximum Likelihood Method).

تطبق هذه الطريقة فقط على النموذج اللوجستي أحادي المعلمة والنماذج المنبقة عنه، حيث يكون اقتران الاحتمالي (Likelihood Function) مشروطاً بعدد الإجابات الصحيحة، إلا أنه لا يمكن استخدام هذا الأسلوب في تقدير المعالم في حالة الإجابة الصحيحة على جميع فقرات الاختبار، أو الإجابة الخاطئة على جميع فقرات الاختبار. ويُؤخذ على هذا الأسلوب أن المفحوصين الذين يحصلون على نفس العدد من الإجابات الصحيحة سيحصلون على نفس تقدير القدرة بالرغم من اختلاف الفقرات التي اجابوا عليها إجابة صحيحة.

ثانياً: طريقة تقدير بيز (Bayesian Estimation Method):

تُستخدم هذه الطريقة عادةً عندما لا يمكن استخدام طرق الأرجحية العظمى، وذلك عندما يُجيب الفرد

على جميع فقرات الاختبار إجابة صحيحة أو إجابة خاطئة، وتمتاز باستخدامها معلومات أولية (Priori Information) متوفرة من خبرات سابقة إضافة إلى استخدامها معلومات من العينة. وهذا يعني انه إذا توافرت معلومات عن توزيع القدرة لدى مجموعة من الافراد المفحوصين، فإن الأساليب التي تعتمد على طريقة "ببيز" يمكن أن تساعد في الحصول على تقديرات جيدة لهذه المعالم.

دالة معلومات الفقرة (Item Information Function).

لكل فقرة من فقرات الاختبار، هنالك ما يسمى بدالة أو منحنى معلومات الفقرة، وهو عبارة عن إقتران يبين مدى مساهمة الفقرة في تقدير القدرة. وبشكل عام فإن الفقرات ذات التمييز العالي تساهم بقوة أكبر في تأكيد دقة قياس القدرة للأفراد من تلك الفقرات ذات التمييز المتدني، حيث تُعطي الفقرة أفضل مساهمة لها في تأكيد دقة قياس القدرة للأفراد عندما تكون صعوبة الفقرة عند مستوى القدرة أو قريبة منها ($\theta=b$). فكلما كان المنحنى مزاحاً باتجاه اليمين فهذا يعني أن الفقرة صعبة، وإذا كان ارتفاع المنحنى عالياً فهذا يعني أن تمييز الفقرة عالٍ. وتختلف معادلة منحنى معلومات الفقرة حسب النماذج اللوجستية (Hambleton & Swaminathan, 1985)، ففي النموذج اللوجستي ثلاثي المعلمة تُعطى دالة معلومات الفقرة بالعلاقة التالية:

$$I_i(\theta) = a_i^2 \left[\frac{Q_i(\theta)}{P_i(\theta)} \right] \left[\frac{P_i(\theta) - C_i^2}{(1 - C_i)^2} \right] \dots \dots \dots (4)$$

$$P_i(\theta) = C_i + (1 - C_i) / (1 + \exp(-L)) \quad \text{حيثُ } P_i(\theta) \text{ تساوي:}$$

$$L = \theta - b \quad \text{L تساوي:}$$

$$Q_i(\theta) = 1 - P_i(\theta) \quad Q_i(\theta) \text{ تساوي:}$$

دالة معلومات الاختبار (Test Information Function).

لقد أشار هاملتون وسواميناثان (Hambleton & Swaminathan, 1985) إلى دالة معلومات الاختبار بالعلاقة التالية:

$$I(\theta) = \sum_{i=1}^N I_i(\theta) \dots \dots \dots (5)$$

حيثُ:

$I_i(\theta)$: كمية المعلومات التي تقدمها الفقرة (i) عند مستوى القدرة (θ).

N: عدد الفقرات في الاختبار.

$I(\theta)$: كمية المعلومات التي يقدمها الاختبار عند مستوى القدرة (θ).

إن دالة معلومات الاختبار ناتجة عن جمع كمية المعلومات التي تُقدمها الفقرات عند مستوى قدرة محدد، وإن زيادة عدد الفقرات تقلل الخطأ المعياري. وبما أن مربع الخطأ المعياري في التقدير يرتبط بالثبات، فإن ذلك يشير إلى أنه إذا زادت كمية معلومات الاختبار فإن مقدار الخطأ المعياري سوف يقل مما يؤدي إلى زيادة الثبات؛ لذلك ترتبط دالة معلومات الاختبار بالثبات ودقة القياس.

وقد بين هاملتون وسواميناثان وروجرز (Hambleton & Swaminathan & Rogers, 1991)

أنه للحصول على تقديرات المعالم بدقة عالية يجب الموازنة بين صعوبة الفقرة من جهة، وقدرات المفحوصين من جهة أخرى، ولن معالم الفقرات السهلة لا يتم تقديرها بدقة عالية في المجموعات ذات القدرات العالية، ومعالم الفقرات الصعبة لا يتم تقديرها بدقة عالية في المجموعات ذات القدرات المتدنية.

وقد أشار بيكر (Baker, 2001) أن منحني معلومات الاختبار يُقدم فوائد عديدة، إذ يُقدم معلومات عن دقة الاختبار في تقدير قدرات الأفراد، وأشار إلى أن الوضع المثالي يتطلب أن يكون هذا المنحنى أفقياً، إلا أن هذا الوضع المثالي لا يتناسب مع أغراض بعض الاختبارات. وحيث أن شكل منحني دالة معلومات الاختبار عادةً يشبه شكل منحني معلومات الفقرة، حيث تكون عادةً جرسية الشكل وتقدر درجات الأفراد بدرجات متفاوتة من الدقة.

الأداء التفاضلي للفقرة ودالة معلومات الفقرة:

في نظرية الاستجابة للفقرة (IRT) يتم عادة تصميم الاختبارات وتطويرها حسب دالة المعلومات، حيث تكون معلومات الفقرة إحدى المقاييس لقياس الدقة في تقدير قدرات الأفراد بناءً على المعلومات التي تقدمها الفقرة، ويتم استخدام دالة المعلومات من أجل اختيار الأسئلة التي تميز بطريقة دقيقة ومرغوبة من الناحية السيكمترية عند مواقع محددة عند أي مستوى من مستويات القدرة. ولأن دالة معلومات الاختبار ناتجة عن جمع كمية المعلومات التي تُقدمها الفقرات عند مستوى قدرة محدد، فإن كمية المعلومات تلعب على مستوى الفقرة دوراً رئيسياً في تطوير الاختبار، وهذا يبدو واضحاً بشكل خاص في الاختبارات المحكية واختبارات التكيف الحاسوبي (Samejima, 1977).

ومن جانب آخر؛ فإن إحدى الجوانب المهمة الأخرى في تطبيق الاختبار هو الكشف عن الأداء التفاضلي للفقرة، والذي من المعروف أنه يظهر حين تكون هناك فروق في أداءات استجابة الفقرة بين المجموعات الفرعية من المفحوصين قيد الدراسة والبحث (Lord, 1980; Raju, 1988).

ويسبب أهمية الأداء التفاضلي للفقرة للكشف عن الفقرات التي تؤثر على احتمالية الاستجابة الصحيحة للمفحوص، وبسبب تطبيقات معلومات الفقرة في إختيار الأسئلة التي يتم وضعها في نماذج الاختبار؛ فإن الدراسة الحالية كان من إحدى أهدافها محاولة الكشف عن العلاقة بين الأداء التفاضلي للفقرة وبين معلومات الفقرة.

ثانياً: التحيز والأداء التفاضلي للفقرة: (Bias and Differential Item Functioning).

لقد نظر الباحثون لمفهوم التحيز على أنه مصطلح يخفي وراءه معنى سلبياً ويرتبط بشكل عام بفكرة عدم العدالة، والاهتمام بالمفاهيم المشحونة بالجانب السياسي والاجتماعي أكثر من ارتباطه بالجانب السيكومتري. لذا عمل المتخصصون في مجال بناء الاختبارات النفسية على استخدام مفهوم الأداء التفاضلي للفقرة منذ بداية الثمانينات للتعبير عن الطرق والمعاني الإحصائية المستخدمة للكشف عن طرق تحيز فقرات الاختبار من منطلق أنه من الأفضل استخدام المصطلحات الفنية السيكومترية بدلاً من المصطلحات المشحونة سياسياً واجتماعياً (Raju & Ellis, 2003).

ويحدث أداء الفقرة التفاضلي نتيجةً لبعض الأخطاء المنهجية الناتجة عن معالم الفقرة ومعالم القدرات للأفراد والتي يتم تقديرها (Hambleton & Swaminathan & Rogers, 1991). ويمكن توضيح ذلك من خلال القول بأن الاختبار متعدد الأبعاد، ومن الناحية الأساسية لا يمكن القول بأن هناك اختبار أحادي البعد بشكل كامل على الرغم أنه يمكن تطوير اختبار أحادي البعد بشكل جيد. وهناك دائماً أبعاد ثانوية في الاختبار، وتقدم الأبعاد الثانوية غير المستهدفة أخطاءً منهجية في البيانات والتي تعمل على التأثير بشكل سلبي على الفروقات الجماعية على البعد المستهدف. ولهذا السبب، فإن الباحث الذي يطبق منحنى خصائص فقرة أحادي البعد على فقرة تتطلب قدرتين مختلفتين، فإن معرفة أو عدم معرفة القدرات غير

المستهدفة يمكن أن يؤثر على منحى خصائص الفقرة بالنسبة للمجموعة. مع ذلك؛ فإن تعددية البعد لا تؤدي بالضرورة إلى الأداء التفاضلي للفقرة. وإن القدرة غير المتساوية على الأبعاد الثانوية هي التي تسبب وجود الأداء التفاضلي للفقرة. ويجب الإشارة إلى أن هذه الأبعاد الثانوية يطلق عليها في بعض الأحيان الأبعاد الثانوية الصغيرة في الأدب السابق.

الأداء التفاضلي للفقرة: (Differential Item Functioning: DIF).

إن الأداء التفاضلي للفقرة دلالة مشتقة إحصائياً للتعبير عن الفروق في احتمالات الاستجابة الصحيحة للفقرة بين مجموعتين في نفس مستوى القدرة (Jodoin & Geirl & Ackerman, 2000; Teresi, 2001)، والأداء التفاضلي للفقرة من منظور نظرية الاستجابة للفقرة يتم من خلال فحص العلاقة بين الاحتمالات المشروطة للاستجابات الصحيحة على الفقرة وبين السمة الكامنة المراد قياسها (θ)، بحيث يتم مقارنة منحنيات خصائص الفقرة (ICCS) في المجموعتين الفرعيتين لمجتمع إحصائي لكشف الأداء التفاضلي، والتي يتم تمثيلها وفق النموذج اللوجستي المستخدم .

ومن الجدير بالذكر أن هنالك مصطلح الأداء التفاضلي للاختبار (Differential Test Functioning: DTF)، وهو خطأ منتظم في القياس ويؤثر على صدق الاختبار أكثر من ثباته (Osterlind, 1983)، والذي يشير إلى الفارق في الأداء على الاختبار، بين مجموعتين مختلفتين من حيث العرق أو الجنس أو مكان السكن أو المستوى الاقتصادي والاجتماعي. ويمكن التحقق من الفارق في الأداء باستخدام طرق تعتمد على النظرية الحديثة في القياس مثل الفرق بين منحنيات خصائص الاختبار (Test Characteristics

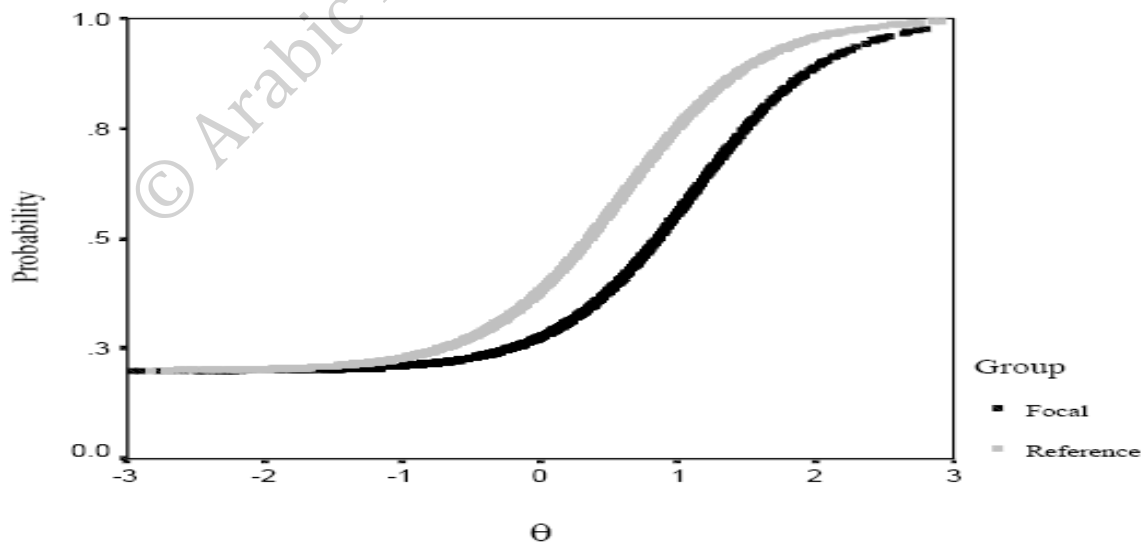
(Curve) في المجموعات المختلفة. والأداء التفاضلي للاختبار أهم بكثير من الأداء التفاضلي للفقرة، لأن القرارات التي تتخذ تعتمد على علامات الاختبار ككل وليس على علامات الفقرات المفردة.

أنواع الأداء التفاضلي:

بناءً على العديد من نتائج الدراسات (Hambleton & Rogers, 1989; pae, 2004):

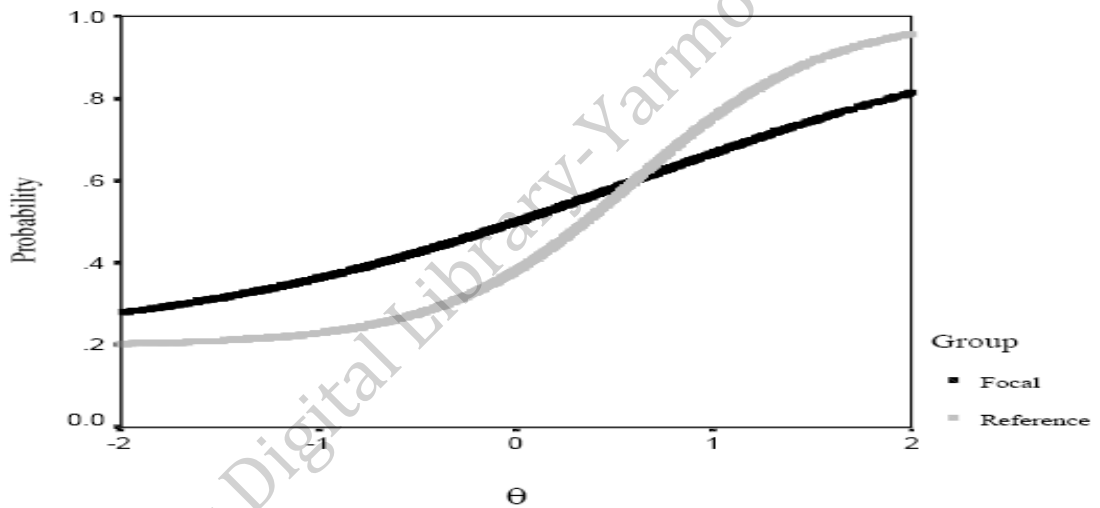
(Swaminathan & Rogers, 1990) ظهر نوعان من الأداء التفاضلي هما:

1. الأداء التفاضلي المنتظم (Uniform DIF): ينتج الأداء التفاضلي المنتظم لفقرة عندما لا يحدث تفاعل بين مستوى القدرة والمجموعة التي ينتمي إليها الفرد، أي أن احتمال الإجابة الصحيحة للفقرة يكون أكبر دائماً لإحدى المجموعتين عند جميع مستويات القدرة. حيث تكون معالم التمييز والتخمين متساوية بين المجموعات، ولكن تكون معلمة الصعوبة لمنحنى خصائص الفقرة غير متساوية بين المجموعات، والشكل (2) يوضح ذلك.



الشكل (2): منحنى الفقرة الذي يظهر أداءً تفاضلياً منتظماً بين المجموعات.

2. الأداء التفاضلي غير المنتظم (Nonuniform DIF): يظهر عندما يكون هنالك تفاعل بين مستوى القدرة والمجموعة التي ينتمي إليها الفرد، فنمط الفروق في احتمالية الاستجابة الصحيحة ليست متشابهة في كل مستويات القدرة، فهذه الفروق قد تكون لصالح إحدى المجموعات عند مستوى قدرة معين، وعند مستوى قدرة آخر يظهر لصالح المجموعة الأخرى. حيث تكون معلمة التمييز مختلفة عبر المجموعات ومعلمة الصعوبة والتخمين متساوية، والشكل (3) يوضح ذلك.



الشكل (3): منحني الفقرة الذي يظهر أداءً تفاضلياً غير منتظم بين المجموعات.

طرق الكشف عن الأداء التفاضلي للفقرة:

يوجد العديد من الطرق للكشف عن الأداء التفاضلي للفقرة، فبعضها يعتمد على النظرية الكلاسيكية في القياس، مثل طريقة تحليل التباين (Analysis of variance)، وطريقة مربع كاي (Chi-Square, χ^2)، وطريقة مانتل هانزل (Method Mantel-Haenzel)، وبعضها الآخر يستند إلى النظرية الحديثة في القياس، مثل طريقة فرق المساحة لمنحني خصائص الفقرة (Item Characteristic

(Curve Methods, ICC)، وطريقة نسبة الأرجحية: (IRT Likelihood - Ratio Tests For DIF). (IRTLRDIF)، حيث تم في هذه الدراسة الاكتفاء بعرض طريقة فرق المساحة لمنحنى خصائص الفقرة بقدر مناسب من التوضيح لأنها تستند إلى النظرية الحديثة في القياس.

طريقة فرق المساحة بين منحنيات خصائص الفقرة (ICC).

تعتمد الطرق المتبعة في الكشف عن الأداء التفاضلي للفقرة وفق نماذج منحنى خصائص الفقرة على التشابه في منحنيات خصائص الفقرة للمجموعات الفرعية المختلفة، فإذا تشابهت منحنيات خصائص الفقرة في المجموعات الفرعية المختلفة فإن الفقرة لا تظهر أداءً تفاضلياً، أي تكون الفقرات متساوية في صعوبتها وتمييزها وتخمينها لجميع المفحوصين من نفس مستوى القدرة في المجموعات الفرعية المختلفة، وذلك حسب النموذج المستخدم (Chung & Huisu, 2004; Crocker & Algina, 1986).

ويصف كل من هاملتون وروجرز (Hambleton & Rogers, 1995) إجراءات الكشف عن الأداء التفاضلي للفقرة في ثلاث فئات من الطرق وهي: الطرق القائمة على مقارنة منحنيات خصائص الفقرة (ICC)، والطرق القائمة على مقارنة مطابقة نماذج استجابة الفقرة للبيانات في المجموعات المختلفة للمفحوصين، والطرق القائمة على مقارنة معالم الفقرة. ويمثل النموذج اللوجستي ثلاثي المعلم الشكل العام للكشف عن الأداء التفاضلي للفقرة، لأنه يضم المعالم الثلاث الممكنة للفقرة (a,b,c).

وقد بين راجو (Rajua, 1988) أنه يمكن حساب المساحة بين المنحنيات لخصائص الفقرة، سواء ذات المعلمة الواحدة أو المعلمتين أو الثلاثة معالم. ففي النموذج ثلاثي المعالم يُعبر عن المساحة بالمعادلة التالية:

$$Area = (1 - c) \left| \frac{2(a_2 - a_1)}{D a_1 a_2} \ln \left(1 + \exp \left(\frac{D a_1 a_2 (b_2 - b_1)}{a_2 - a_1} \right) \right) - (b_2 - b_1) \right| \dots \dots (6)$$

وعند استخدام النموذج ثنائي المعلمة فإنه لا يوجد تخمين (مقدار C) وتصبح المعادلة:

$$Area = \left| \frac{2(a_2 - a_1)}{D a_1 a_2} \ln \left(1 + \exp \left(\frac{D a_1 a_2 (b_2 - b_1)}{a_2 - a_1} \right) \right) - (b_2 - b_1) \right| \dots \dots (7)$$

حيث يمثل الرمز (C) معلمة التخمين، وتم حسابها بالمتوسط الحسابي لمعلمتي التخمين في المجموعتين، ويمثل الرمز (a_1) معلمة التمييز للمجموعة الأولى، و (a_2) معلمة التمييز للمجموعة الثانية، ويمثل الرمز (b_1) معلمة الصعوبة للمجموعة الأولى، و (b_2) معلمة الصعوبة للمجموعة الثانية، والرمز (D) الذي يشير إلى قيمة ثابتة وتساوي (1.7). وعند استخدام النموذج اللوجستي أحادي المعلمة فإن المساحة تكون القيمة المطلقة للفرق بين معلمة الصعوبة للمجموعتين.

$$AREA = |(b_2 - b_1)| \dots \dots (8)$$

وقد اقترح زيكي (Zieky) المشار إليه في دريانا (Driana, 2007) باستخدامه طريقة "مانتل هانزل" مقاييس لحجم الأثر، والتي يمكن استخدامها من أجل تكيم مستوى الأداء التفاضلي للفقرة. وقد تم استخدام هذه القاعدة أيضاً من أجل تكيم مستوى الأداء التفاضلي ضمن سياق نظرية الاستجابة للفقرة على الشكل التالي:

1. مستوى (A) من الفقرات ذات الأداء التفاضلي، وهي الفقرات ذات الأداء التفاضلي المهم.
2. مستوى (B) من الفقرات ذات الأداء التفاضلي، وهي الفقرات ذات الأداء التفاضلي المتوسط، إذا كانت

$$\Delta_{aMH} \text{ دالة احصائياً عند مستوى الدلالة } 0.05 \text{ و } |\Delta_{aMH}| \geq 1 \text{ و } 1.5 > |\Delta_{aMH}|$$

3. مستوى (C) من الفقرات ذات الأداء التفاضلي: أداء تفاضلي عالٍ، إذا كانت قيمة Δ_{aMH} دالة إحصائياً

عند مستوى الدلالة 0.05 و $|\Delta_{aMH}| \leq 1.5$.

ولقد أشار لين وهارنيس (Linn & Harnisch, 1981) إلى طريقة أخرى للكشف عن الأداء التفاضلي للفقرة تقوم على مطابقة نموذج الاستجابة للفقرة سواء للنموذج الأحادي أو الثنائي أو الثلاثي المعلمة للمجموعة المستهدفة مع نموذج استجابة الفقرة للمجموعة المرجعية، حيث يتم في هذه الطريقة حساب معالم الفقرة ومعالم القدرة في كل من المجموعتين المرجعية والمستهدفة. ويستدل على الأداء التفاضلي للفقرة إذا كانت معالم الفقرة للمجموعة المرجعية لا تطابق معالم الفقرة للمجموعة المستهدفة، وإذا تطابقت معالم الفقرة للمجموعة المرجعية مع معالم الفقرة للمجموعة المستهدفة فإن ذلك يشير إلى عدم وجود أداء تفاضلي للفقرة. ويمكن اختبار الدلالة الإحصائية لجودة المطابقة باستخدام الإحصائي كاي تربيع (χ^2). ومن سلبيات هذه الطريقة أن المجموعة المستهدفة قد تكون قليلة العدد فلا يكون تقدير معالم الفقرة لها دقيقاً.

وبغض النظر عن المؤشرات الإحصائية المستخدمة للكشف عن الأداء التفاضلي للفقرة وفقاً لنموذج

الاستجابة للفقرة المستخدم؛ فإنه يجب إتباع الخطوات التالية (Ho Kim & Cohen, 1995):

1. التحقق من افتراض أحادية البعد (Unidimensionality) في الاختبار المستخدم لقياس السمة

أو القدرة موضع الاهتمام.

2. وضع تقديرات معالم الفقرة للبيانات المتحققة في المجموعات الفرعية المختلفة على مقياس

موحد (Same Metric) قبل مقارنة منحنيات خصائص الفقرة للمجموعات.

3. تحديد النموذج اللوجستي المراد استخدامه لمقارنة المنحنيات.

4. تحديد الطرق أو المؤشرات التي ستستخدم للكشف عن التحيز.

الدراسات السابقة

بعد مراجعة الأدب التربوي والإطلاع على الدراسات التي اهتمت بدراسة الأداء التفاضلي تبين أن هنالك اهتمام كبير في الأدب النظري والتجريبي للكشف عن الأداء التفاضلي للفقرة في الاختبارات والمقاييس النفسية والتربوية، وقد أخذ هذا الاهتمام عدة جوانب فمنها ما حاول المقارنة بين فاعلية الطرق في الكشف عن الأداء التفاضلي للفقرة، خاصة مقارنة الطرق التي تعتمد نظرياً على نظرية الاستجابة للفقرة، ومنها ما حاول دراسة الظروف التي تؤثر على فاعلية الطرق في الكشف عن الأداء التفاضلي للفقرة. وهنالك دراسات قليلة جداً حاولت دراسة الأثر الذي ينتج من وجود الفقرات ذات أداء تفاضلي في الاختبار في ضوء بعض المتغيرات.

وتتنوع الدراسات من حيث طبيعة البيانات، فاستخدم بعضها بيانات حقيقية مأخوذة من اختبارات مقننة، واستخدم بعضها الآخر بيانات مولدة (simulated data) باستخدام برمجيات متخصصة. ومن هذه البرمجيات ما يمكنه توليد بيانات بمعالم مختلفة، ومطابقة لنماذج مختلفة، ومنها ما يولد بيانات مطابقة لبيانات اختبارات ومقاييس مقننة. وقد لجأ الدارسون إلى البيانات المولدة ليتمكنوا من تغيير الظروف التي من الممكن أن يؤثر فيها الأداء التفاضلي للفقرة في ضوء بعض المتغيرات، وكذلك لقدرتهم على فحص فاعلية طرق الكشف عن الأداء التفاضلي للفقرات التي تبدي أداء تفاضلياً.

ولخدمة أغراض الدراسة، وسهولة استعراض الدراسات السابقة؛ فقد تم تصنيفها إلى صنفين:

- 1- الدراسات التي اهتمت بدراسة أثر الفقرات ذات الأداء التفاضلي في الاختبار في بعض المتغيرات، مثل معادلة الاختبار وإحصائيات النظرية الكلاسيكية (الخطأ المعياري في القياس، ومعامل الثبات، ومعامل الارتباط المصحح، ومعامل كرونباخ الفا)، ومستوى الخطأ من النوع الأول وحجم الأثر لاختبار ت

(t-test) للعلامات الملاحظة الكلية.

2- الدراسات التي قارنت فاعلية طرق الكشف عن الأداء التفاضلي للفقرة، وخاصة تلك الطرق التي تعتمد على نظرية الاستجابة للفقرة باختلاف بعض الظروف، مثل حجم العينة وعدد الفقرات ونوع الأداء التفاضلي (منتظم-غير منتظم) ومصادر الأداء التفاضلي.

أولاً: الدراسات التي اهتمت بدراسة أثر الأداء التفاضلي للفقرات الموجودة في الاختبار في ضوء بعض المتغيرات.

أجرى زومبو (Zumbo, 2000) دراسة هدفت للكشف عن أثر الفقرات ذات الأداء التفاضلي على إحصائيات نظرية الاختبار الكلاسيكية، وهي: معاملات الارتباط، ومعامل كرونباخ ألفا، والخطأ المعياري في القياس، ومعامل الارتباط المصحح لأثر الإنخفاض. واستخدم في هذه الدراسة نتائج اختبار التوفل (TOFEL)، وذلك باستخدام النموذج اللوجستي ثلاثي المعلمة في ظل مجموعة متنوعة من ظروف الاختبار. وقد كان حجم الاختبار ثابتاً مقداره (38) فقرة. وكانت عدد الفقرات التي تبدي أداءً تفاضلياً هي فقرة واحدة، وكان مستوى الأداء التفاضلي للفقرة (0.2, 0.5, 1.0)، وهذا يعتبر أحد الظروف التجريبية المحبذة من أجل الكشف فيما إذا كان هناك فقرة واحدة من الفقرات ذات الأداء التفاضلي قادرة على التأثير في إحصائيات الاختبار في ضوء النظرية الكلاسيكية في القياس. وركزت الدراسة على نوع واحد من الأداء التفاضلي وهو الأداء التفاضلي المنتظم، وهذا يعكس أحد المواقف التي يمكن من خلالها عزو وجود أداء الفقرة التفاضلي في الفروقات في صعوبة الفقرة فقط. وأشارت النتائج إلى أن أداء الفقرة التفاضلي لم يكن ذا أثر دال إحصائياً على إحصائيات الاختبار في ضوء النظرية الكلاسيكية في القياس.

وقام إنجل وقيدون وجوز (Angel & Gideon & Jose, 2000) بدراسة هدفت للكشف عن أثر نسبة الفقرات ذات الأداء التفاضلي في الاختبار على قوة وفاعلية طريقة "مانتل هانزل" (MH) في الكشف عن الفقرات ذات الأداء التفاضلي، وقد اعتمدت النسب التالية للفقرات ذات الأداء التفاضلي الموجودة في الاختبار (10%, 20%, 30%)، وثلاثة أطوال للاختبار (20, 40, 60) فقرة. وتم توليد بيانات الاستجابة على الفقرات باستخدام النموذج اللوجستي ثلاثي المعلمة (3PLM) للمجموعات المرجعية والمستهدفة. وأشارت النتائج أنه عندما تكون الفقرات ذات الأداء التفاضلي تشكل نسبة (10%) من فقرات الاختبار فإن إجراءات طريقة مانتل هانزل (MH) تتصف بالفعالية وبالقوة الإحصائية الكافية للكشف عن الأداء التفاضلي للفقرة، ولكن حين يكون عدد الفقرات ذات الأداء التفاضلي تشكل ما نسبته (15%-30%) فإن قوة وفعالية طريقة مانتل هانزل ستخفض، كما أشارت النتائج أيضاً إلى أن أثر طول الاختبار على القوة الإحصائية لمؤشر "مانتل هانزل" كان منخفضاً.

كما أجرى جون ويليم وبربارا (John & William & Barbara, 2002) دراسة هدفت للكشف عن تأثير الفقرات ذات الأداء التفاضلي على تقدير الفروقات العمرية والجنسية في مستوى شدة الإعاقة لدى عينات من الأفراد الراشدين الذين يعانون من الإعاقة الوظيفية. وجمعت البيانات من خلال عينة مكونة من (750) من المستجيبين الراشدين الذين يحصلون على المساعدة أو على الإشراف في (11) نشاطاً من النشاطات اليومية التي يقومون بها. وتم في هذه الدراسة تقدير الفروق الجنسية والفروق العمرية من حيث الإعاقة وذلك باستخدام نماذج متعددة. أشارت النتائج إلى أن (9) فقرات أظهرت أداءً تفاضلياً في ضوء الفروق العمرية والجنسية. وكان الأداء التفاضلي للفقرة واضحاً بشكل خاص في الفقرات التي تقيس نشاطات التسوق وإدارة الأمور. وبدون تعديل أي واحدة من الفقرات ذات الأداء التفاضلي أشارت النتائج إلى أن الأفراد

متوسطي العمر كانوا أقل إعاقة وظيفية مقارنة مع الأفراد المسنين، وأن الإناث كنَّ أقل إعاقة مقارنة مع الذكور. وبعد أن تم تصحيح الفقرات ذات الأداء التفاضلي؛ أشارت النتائج إلى عدم وجود فروق في الإعاقات الوظيفية بين الأفراد المعاقين متوسطي العمر وبين الأفراد المعاقين المسنين، ولم تكن هناك فروق دالة إحصائية تُعزى إلى الجنس.

وقام لوبيسكو (Luppescu, 2002) بدراسة هدفت الى المقارنة بين النماذج الخطية الهرمية للكشف عن الأداء التفاضلي للفقرة مقارنة مع الطرق التقليدية مثل مؤشر فرق معلمة الصعوبة (b) وفقاً لنموذج راش، وهو من مؤشرات النظرية الحديثة في القياس. وميزة هذه النماذج أنها قادرة على دقة تقدير القدرة للفرد بوجود الفقرات ذات الأداء التفاضلي. وتم في هذه الدراسة توليد ثلاثة أحجام للعينات هي (500,250,100)، وأربعة مستويات لدرجة الأداء التفاضلي، وهي (0.25 , 0.50 , 0.75 , 1.00)، وثلاث نسب للفقرات ذات الأداء التفاضلي الموجودة في الاختبار، وهي (10%, 25%, 50%). أشارت النتائج في معظم الحالات إلى تساوي قيم الجذر التربيعي لمتوسط مربعات الخطأ (RESE) لتقدير صعوبة الفقرة ومعالم القدرة لنموذج راش والنماذج الخطية الهرمية (HLM). واستنتج من ذلك الحالات التي كان فيها حجم العينة صغيراً، إذ كانت قيم (RMSE) أكبر في النماذج الخطية الهرمية، وفي حال ازدياد حجم العينة قل الأداء التفاضلي للفقرات، وخلصت الدراسة إلى أن النماذج الخطية الهرمية أعطت تقديرات أعلى من الطرق التقليدية.

وأجرى زومبو (Zumbo, 2003) دراسة هدفت للكشف فيما إذا كان أداء الفقرة التفاضلي على مستوى الفقرة يظهر من خلال التحليلات على مستوى المقياس الكلي مثل استخدام التحليل العاملي متعدد المجموعات، والأحادي، وحساب معامل ألفا لكل مجموعة. وتم تقديم بيانات خاصة لدراسة توليد البيانات

باستخدام معالم الفقرة التي تم الحصول عليها من اختبار اللغة الإنجليزية توفل (TOFEL) نتيجة لترجمة هذه الاختبارات من لغة إلى لغة أخرى من أجل الكشف فيما إذا كان الأداء التفاضلي للفقرة على مستوى الفقرة يظهر من خلال التحليل على مستوى المقياس. وولدت بيانات استجابات المفحوصين باستخدام برنامج (MIRTGN)، وذلك باستخدام نموذج الاستجابة للفقرة اللوجستي ثلاثي المعلمة تحت ظروف مختلفة من الاختبارات، وهي: الفروق في مستويات الأداء التفاضلي للفقرة، والتي تراوحت بين مستويات متوسطة إلى مستويات مرتفعة، واستخدام الأداء التفاضلي المنتظم فقط، كما واستخدمت أعداد مختلفة من الفقرات ذات الأداء التفاضلي، وهي (16 , 8 , 4 , 1) فقرة من مجموع الفقرات الكلي الذي بلغ (38) فقرة.

أشارت النتائج أنه بغض النظر عن مستوى الأداء التفاضلي للفقرة، ونسبة الفقرات ذات الأداء التفاضلي في الاختبار، فقد أشار التحليل العاملي التوكيدي للمجموعات المتعددة أن التشبعات العاملية وتباينات الخطأ كانت متطابقة من الناحية الإحصائية للاختبار الأصلي والاختبار المترجم على الرغم أن الاختبار المترجم قد احتوى على ما نسبته (24.1%) من الفقرات التي تظهر مستويات عالية من الأداء التفاضلي للفقرة، أي أن الأداء التفاضلي للفقرة على مستوى الفقرة لم يظهر في النتائج على مستوى مقياس التدرج، وبالتالي فإنه من الضروري عند محاولات ترجمة الاختبارات اللغوية أن يؤخذ بعين الاعتبار أداء الفقرة التفاضلي على مستوى الفقرة، وأن لا يتم الإعتماد على الكشف عن أداء الفقرة التفاضلي على مستوى المقياس الكلي.

وأجرى كاماتا وشو (kamata & chu, 2003) دراسة هدفت إلى دقة تقديرات معالم الصعوبة للفقرات ومعالم القدرة للأفراد ومعادلة درجات الاختبار باستخدام نموذج راش الهرمي بوجود فقرات ذات أداء تفاضلي. وتم تخصيص مجموعتين عشوائيتين من المفحوصين لكل نموذج، وتم توليد (500) نمط استجابة لكل مجموعة من المجموعتين وكل مجموعة من المجموعات طبق عليها نموذج مختلف من الاختبار وتكون كل

نموذج من (20) فقرة كان من بينها (10) فقرات مشتركة. حيث تم تقييم فاعلية أداء نموذج راش الهرمي في حال وجود فقرات مشتركة ذات أداء تفاضلي وفي حال حذفها. كما تم أيضاً اختبار عدد الفقرات ذات الأداء التفاضلي وكانت (10,5,2)، ومستوى الأداء التفاضلي حسب مؤشر ماننل هانزل (06, 0.8, 1.2, 2.00). واجري بعد ذلك تنفيذ المعادلة المتزامنة وتصميم المجموعات العشوائية غير المتكافئة ذات الاختبار المشترك لمقارنتهما مع نموذج راش الهرمي، واستخدمت برمجية (HLM) لتقدير المعالم (معلمة الصعوبة ومعلمة القدرة) في نموذج راش الهرمي، واستخدم برنامج (Bilog-Mg3) لتقدير المعالم (معلمة الصعوبة ومعلمة القدرة) في النموذج الأساسي (راش)، واستخدم في الدراسة المعايير التالية لتقييم دقة المعادلة وهي: الجذر التربيعي لمتوسطات مربعات خطأ المعادلة (RMSE)، ومربع التمييز والخطأ العشوائي.

أشارت النتائج إلى أن فاعلية نموذج راش الهرمي أفضل من نموذج راش التقليدي فيما يتعلق بدقة تقدير معالم الصعوبة للفقرات وقدرات الأفراد بوجود الفقرات ذات الأداء التفاضلي، فقد كان متوسط مربع الخطأ لتقدير معالم الصعوبة لفقرات الاختبار باستخدام نموذج راش الهرمي أقل من النموذج الأساسي (راش) الذي لا يحتوي على فقرات ذات أداء تفاضلي، وتأثر نموذج راش (النموذج الأساسي) بشكل كبير بعدد الفقرات ذات الأداء التفاضلي، وكذلك بمستوى الأداء التفاضلي للفقرة.

كما أجرى زهانغ وماثيوز ودورانز (Zhang & Matthews & Dorans, 2003) دراسة هدفت لتحديد أثر الفقرات ذات الأداء التفاضلي في المعادلة باستخدام اختبار سات (SAT) وهو اختبار الاستعداد المدرسي، فقد كان مستوى الأداء التفاضلي لهذه الفقرات عالياً ($DIF < 1.5$). وتم استخدام طريقة المعادلة المئينية، وتصميم المجموعة الواحدة وتم تدرج العلامات إستناداً لنموذج الاختبار القبلي المؤلف من (35) فقرة، وإجراء معادلة له مع الاختبار الأساسي المكون من (76) فقرة. حيث تم الكشف عن (3) فقرات ذات

أداء تفاضلي عالٍ باستخدام مؤشر "مانتل هانزل" (MH)، وتم إجراء المعادلة قبل حذف الفقرات ذات الأداء التفاضلي في الاختبار القبلي مع الاختبار الأساسي، وتم أيضاً إجراء المعادلة بعد حذف الفقرات ذات الأداء التفاضلي في الاختبار القبلي مع الاختبار الأساسي. أشارت النتائج إلى أن حذف الفقرات ذات الأداء التفاضلي كان لصالح المجموعة المرجعية، في حين تأثر تدرج المجموعات المستهدفة سلبياً بحذف الفقرات ذات الأداء التفاضلي.

وفي دراسة أجراها زهان وزومبو (Zhan & Zumbo, 2009) هدفت للكشف عن أثر الفقرات ذات الأداء التفاضلي على الخطأ من النوع الأول وحجم الأثر للعينات المستقلة لاختبار (T-Test) من خلال علامات الاختبار الكلية الملاحظة. واستخدم النموذج اللوجستي ثلاثي المعلمة لتقدير المعالم بوجود ثلاثة مستويات للأداء التفاضلي للفقرة (مستوى متدنٍ، مستوى متوسط، مستوى عالٍ) بناءً على الإحصاءات التي قدمها راجو (Raju, 1988). وكانت مستويات الأداء التفاضلي للفقرة على الشكل التالي: (0.4, 0.6, 0.8). كما استخدم أيضاً أربع حالات من عدد الفقرات ذات الأداء التفاضلي في الاختبار، وكانت على الشكل التالي: (1, 4, 8, 16) من عدد الفقرات الكلية في الاختبار والتي بلغت (38) فقرة. وتم استخدام أربع مستويات من حجم العينة هي: (25, 50, 125, 250) من المفحوصين في كل مجموعة. إضافة لذلك، ولأغراض المقارنة؛ فقد تم دراسة أن عدم وجود الفقرات ذات الأداء التفاضلي يمثل بيانات القاعدة الخطية. نتيجة لهذه الظروف فقد كانت تجربة توليد البيانات الناتجة على الشكل التالي (4*3) وهو عبارة عن تصميم عاملي متقاطع بشكل كامل.

أشارت النتائج حين يكون لدينا فقرة واحدة فقط من الفقرات ذات الأداء التفاضلي في الاختبار الذي يتكون من (38) فقرة فإن مستوى الخطأ من النوع الأول لم يتأثر باستثناء معيار "راجو" في منطقة (0.8)

حين يكون هنالك (250) مفحوصاً في كل مجموعة. وبغض النظر عن مستوى الأداء التفاضلي للفقرة وحجم العينة حين يكون هنالك (16) فقرة من الفقرات ذات الأداء التفاضلي من مجموع الفقرات الكلي البالغة (38) فقرة حصل تضخم في مستوى الخطأ من النوع الأول. كما أشارت النتائج أيضاً عند وجود الفقرات ذات الأداء في الاختبار تزداد الفروقات في حجم الأثر، فكلما ازداد عدد فقرات ذات الأداء التفاضلي في الاختبار وزاد مستوى الأداء التفاضلي للفقرة ازدادت الفروقات في حجم الأثر، حيث تكون الفروق في المتوسط الحسابي للعلامة الملاحظة متضخمة نتيجة لوجود فقرات ذات الأداء التفاضلي .

أجرى إرحيل (2011) دراسة هدفت للكشف عن أثر الفقرات ذات الأداء التفاضلي للجنس على دقة المعادلة العمودية لاختبار أوتيس لينون للقدرة العقلية بمستوياته المتوسط والمتقدم. تألفت عينة الدراسة من (1001) طالب وطالبة من طلبة الصف السابع والثامن والتاسع مقسمين إلى (500) طالب في مستوى القدرة المتوسط، و(501) طالب في مستوى القدرة المتقدم. وتم استخدام تصميم الفقرات المشتركة والمجموعات غير المتكافئة، وطبق اختبار المستوى المتوسط على أفراد مستوى القدرة المتوسطة، وطبق اختبار المستوى المتقدم على أفراد القدرة المتقدمة، وتضمن كل من الاختبارين فقرات مشتركة بلغ عددها (23). واستخدم في هذه الدراسة طريقة "مانتل هانزل" (MH) للكشف عن الفقرات ذات الأداء التفاضلي في كل مستوى من مستويات الاختبار بشكل منفصل. وبعد ذلك تم إجراء المعادلة العمودية باستخدام برمجية (Bilog-MG) وفقاً للنموذج أحادي المعلمة، وأجريت المعادلة في ضوء الإحتفاظ بالفقرات ذات الأداء التفاضلي وفي ظل حذف الفقرات ذات الأداء التفاضلي. أشارت النتائج إلى وجود أثر للفقرات ذات الأداء التفاضلي على دقة المعادلة العمودية.

وأجرى الحياصات (2011) دراسة هدفت لتحديد أثر الفقرات ذات الأداء التفاضلي في تحقيق معايير الفاعلية في معادلة الاختبار باستخدام النموذج اللوجستي ثلاثي المعلمة. تم استخدام استجابات الطلبة على نموذجي الاختبار الوطني في الرياضيات، وطبق النموذج الأول على (57002) طالباً وطالبة، وطبق النموذج الثاني على (56581) طالباً وطالبة. وتكون نموذج الاختبار من أربعين فقرة بينهما عشر فقرات مشتركة، واستخدم في هذه الدراسة طريقة "تكر" الخطية التي تتبع النظرية الكلاسيكية، وطريقة معادلة العلامات الحقيقية في ضوء النظرية الحديثة، واعتمد في إجراءات المعادلة على تصميم المجموعات غير المتكافئة ذات الاختبار المشترك وتم استخدام بعض المعايير للحكم على دقة المعادلة.

أشارت النتائج إلى أن حذف الفقرات ذات الأداء التفاضلي للجنس يؤدي الى زيادة دقة المعادلة وفعاليتها وفقاً لطريقة "تكر" الخطية التي تتبع النظرية الكلاسيكية. وبينت النتائج أيضاً أن حذف الفقرات ذات الأداء التفاضلي للجنس يؤدي الى زيادة الدقة للمعادلة وخاصة في الطرف العلوي من العلامات وفقاً لطريقة معادلة العلاقات الحقيقية التي تتبع النظرية الحديثة. كما وأظهرت النتائج أن المعادلة باستخدام طريقة "تكر" الخطية أكثر فاعلية من معادلة العلامات الحقيقية قبل حذف الفقرات ذات الأداء التفاضلي وبعده.

ثانياً: الدراسات التي قارنت فاعلية الطرق المختلفة للكشف عن الأداء التفاضلي للفقرة، وخاصة الطرق التي تعتمد على نظرية الاستجابة للفقرة.

أجرى صبكوفياك وسراج وماك وارنسون (Subkoviak & Crag & Mack & Ironson, 1984) دراسة تهدف إلى مقارنة عدد من طرق الكشف عن الأداء التفاضلي هي: طريقة فرق المساحة لمنحنى خصائص الفقرة ذات الثلاثة معالم، وطريقة كأي تربيع وطريقة تحويل صعوبة الفقرة. حيث استخدم اختبار مقنن في القدرة اللغوية يتكون من (40) فقرة، أضيف إليه (10) فقرات تم صياغتها من قبل متخصص في

المفردات اللغوية الدارجة لدى الأفراد السود، واعتبرت هذه الفقرات متحيزة لهم. كما تكونت عينة الدراسة من (1008) فرد من السود، و(1021) من البيض. وبعد جمع البيانات وحساب مؤشر الأداء التفاضلي بناءً على الطرق المستخدمة لاستجابات الأفراد على الأداة وإجراء التحليل المناسب باستخدام معامل الارتباط؛ تبين أن طريقة فرق المساحة لمنحنى خصائص الفقرة ثلاثية المعالم هي الأكثر فاعلية في الكشف عن الفقرات المعلوم تحيزها المسبق، وبناءً على ذلك ينصح الباحثون باستخدام هذه الطريقة عندما يزيد عدد فقرات الاختبار عن (40) فقرة، ويزيد حجم العينة عن (1000) فرد، خاصة وأن هنالك برامج حاسوبية تتيح استخدام هذه الطريقة مثل برنامج (Logiest). كما جاءت طريقة كأي تربيع في المركز الثاني في قدرتها على الكشف عن الفقرات المتحيزة، حيث ارتبطت طريقة كأي تربيع مع طريقة فرق المساحة لمنحنى خصائص الفقرة ذات الثلاث معالم بصورة واضحة، فقد بلغت قيمة معامل الارتباط (0.97)، فيما احتلت طريقة صعوبة الفقرات المرتبة الأخيرة من حيث قدرتها على الكشف عن فقرات معلوم تحيزها المسبق.

كما أجرى جيليان (Gillian, 1989) دراسة هدفت للكشف عن العلاقة الارتباطية بين الأداء التفاضلي للفقرة (DIF) وبين صعوبة الفقرة. واستخدم في هذه الدراسة إحصائيات لبيانات فقرات لفظية وبيانات لفقرات رياضية في آخر (9) دورات باستخدام اختبار سات (SAT). أظهرت نتائج الدراسة بشكل عام وجود علاقة طردية بين صعوبة الفقرة وبين الأداء التفاضلي للفقرة، وكانت العلاقة الارتباطية بين الأداء التفاضلي وصعوبة الفقرة أقوى على المجال اللفظي للاختبار مقارنة مع المجال الرياضي، أي أن الفقرات الأكثر صعوبة قد أظهرت مستويات إيجابية أكبر من الأداء التفاضلي.

وأجرى والسند وروبسون (Walsted & Robson, 1997) دراسة هدفت إلى الكشف عن الفقرات ذات الأداء التفاضلي لمتغير الجنس لفقرات اختبار في الاقتصاد لطلبة المرحلة الثانوية (Test of

(Economic Literact: TEL) مكون من (46) فقرة من نوع الإختيار من متعدد. جمع الباحثان نتائج الاختبار لعينة مكونة من (2019) مفحوصاً من طلبة المرحلة الثانوية، واستخدمت طريقة فرق المساحة بين منحنيات خصائص الفقرة لمجموعتي الذكور والإناث للكشف عن وجود فقرات ذات أداء تفاضلي لمتغير الجنس. واستخدمت البرمجية الحاسوبية (Bilog MG3) لتقدير معالم الفقرات الثلاث. وكشفت هذه الطريقة عن (6) فقرات ذات أداء تفاضلي لصالح الذكور، وعند حذف هذه الفقرات من اختبار (TEL) وإعادة تطبيق الاختبار المكون من (40) فقرة، تبين للباحثين أن الفروق بين الجنسين قلت وبشكل واضح، وكان متوسط علامات الذكور على الاختبار أعلى من متوسط علامات الإناث بمقدار (1.09)، في حين قل الفارق إلى (0.73) بعد حذف الفقرات ذات الأداء التفاضلي. وخلصت الدراسة إلى أن انخفاض الفروق بين الجنسين تعود لعوامل اجتماعية وعوامل تتعلق بالتدريب على آلية الاختبار المطبق وتعليماته.

كما أجرى انجلهارد وغامر (Engelhard & Gamer, 1999) دراسة هدفت للكشف عن الفروق بين أداء الجنسين في اختبار قدرة رياضية، وهو اختبار يتضمن (60) فقرة من نوع الإختيار من متعدد توزعت على النحو الآتي: (14) فقرة من موضوع تحليل البيانات، و(23) فقرة في موضوع الهندسة والقياس، و(11) فقرة في موضوع الحساب، و(12) فقرة في موضوع الجبر. كما تضمن الاختبار (8) فقرات ذات استجابة محددة البناء. طبق الاختبار على (3970) طالباً من طلبة الصف الحادي عشر في ولاية جورجيا الأمريكية، واستخدمت البرمجية الحاسوبية (Facets) للكشف عن الفقرات ذات الأداء التفاضلي لمتغير الجنس. أشارت نتائج الدراسة إلى أن الطالبات أظهرن فرقاً ذا دلالة إحصائية وثبات أعلى على الطلاب في فقرات الإختيار من متعدد في الجبر، بينما أظهر الذكور فرقاً أقل ثباتاً في الفقرات المتعلقة بالهندسة والأعداد

والعمليات الحسابية، فقد كانت متوسطات الدرجات للطلاب أعلى بدرجة دالة إحصائياً من الإناث في فقرتين من ثماني فقرات متعددة التدرج.

وفي دراسة أجراها عبد العزيز (2006) هدفت إلى الكشف عن درجة التوافق بين منحنيات خصائص الفقرة للكشف عن الأداء التفاضلي لمتغير الجنس لاختبار في الرياضيات لطلبة الصف العاشر ثنائي التدرج ومتعدد التدرج. تم بناء اختبار للصف العاشر من خلال جدول المواصفات كما تم تحديد خصائصه السيكمترية، وأشارت النتائج إلى إن فاعلية طرق منحني خصائص الفقرة في الكشف عن الفقرات ثنائية التدرج التي تظهر أو لا تظهر أداء تفضلياً لمتغير الجنس تقارب (84%) لمعيار طريقة مؤشر المساحة للنموذج اللوجستي ثلاثي المعلمة، و(87%) لمعيار طريقة مؤشر المساحة للنموذج اللوجستي ثنائي المعلمة، و(84%) لمعيار طريقة فرق معلم الصعوبة (b) وفقاً لنموذج راش. أما بالنسبة للفقرات ثنائية التدرج (83%) لمعيار فرق معلم الصعوبة (b)، و(84%) لمعيار طريقة مؤشر المساحة للنموذج ثنائي المعلم. ولم تتفق الطرق الثلاث في تحديد الفقرات التي تظهر أداء تفضلياً، إذ لم تظهر بعض الفقرات التي أظهرت أداء تفضلياً بموجب إحدى الطرق أداء تفضلياً بموجب الطرق الأخرى. فقد بلغ عدد الفقرات التي أظهرت أداء تفضلياً لمتغير الجنس بموجب الطرق الثلاث (9) فقرات، وظهر وجود أداء تفضلي متفاوت الدرجة لمتغير الجنس وفقاً لطبيعة المحتوى الرياضي كان معظمه من النوع غير المنتظم للفقرات ثنائية التدرج ومتعدد التدرج وكان لصالح الإناث في عدد من الفقرات الثنائية التدرج وفقاً لطبيعة القدرة الرياضية.

كما أجرى الحمادنة (2007) دراسة هدفت إلى مقارنة الأداء التفاضلي لفقرات اختبارات تحديد الكفاءة اللغوية في اللغة الانجليزية في الجامعات الأردنية تبعاً لمتغير الجنس وفرع الثانوية العامة، وفق طريقتي فرق المساحة بين منحنيات خصائص الفقرة في النموذج اللوجستي ثلاثي المعلم (ICC)، وطريقة

"مانتل-هانزل" (MH) لمعرفة فيما إذا كان الأداء التفاضلي يختلف باختلاف مستوى القدرة، والتعرف على درجة التوافق بين الطريقتين. ولتحقيق غرض الدراسة تم اختيار ثلاثة نماذج من اختبارات تحديد الكفاءة اللغوية في اللغة الانجليزية، وجميعها من نوع الاختيار من متعدد رباعي البدائل، حيث بلغ عدد فقراتها (100,50,50) فقرة في الجامعات الثلاث (جامعة اليرموك وجامعة آل البيت وجامعة العلوم والتكنولوجيا) تقيس في مجملها مهارات مشتركة في اللغة الانجليزية هي: استيعاب المقروء، القواعد والمفردات، والكتابة، وقُدِّم حاسوبياً في بداية كل عام جامعي. كانت عينة الدراسة (1935) مفحوصاً. أظهرت النتائج أن أكثر المهارات التي أظهرت فقراتها أداءً تفضلياً لمتغير الجنس، هما مهارتا القواعد والمفردات وكانت معظمها لصالح الإناث. فقد كان أكثر عدد من الفقرات التي أظهرت أداءً تفضلياً لمتغير الجنس في اختبار جامعة اليرموك وبنسبة (24%) من فقرات الاختبار، ثم تبعه اختبار جامعة العلوم والتكنولوجيا بنسبة (21%)، بينما بلغت النسبة في اختبار جامعة آل البيت (3%) تقريباً.

وفي دراسة قام بها ريسبر (Risper, 2008) هدفت الى الكشف عن أثر أحجام العينات غير المتساوية في المجموعة المرجعية والمجموعة المستهدفة في الكشف عن الأداء التفاضلي للفقرات المحددة مسبقاً باستخدام اختبار (SiB-Test)، وباستخدام بيانات مولدة. حيث تم استخدام أحجام عينات مكونة من (250,500,1000) مفحوص ونسب هذه الأحجام بين المجموعة المرجعية والمستهدفة كانت على النحو الآتي (1:1, 5:1, 10:1)، ومن ثم تم إنشاء (9) خلايا، وقد تم دراسة الخلايا (9) في ظل أن يكون مستوى الأداء التفاضلي للفقرة (عالٍ ، متوسط)، مما نتج (18) خلية تم دراستها، وذلك باستخدام النموذج اللوجستي ثنائي المعلمة، حيث أشارت النتائج أنه عندما تكون أحجام العينة في المجموعة المستهدفة صغيرة فإنه لا يمكن تقدير معالم الفقرات بشكل جيد مما يؤدي إلى عدم الكفاية في القوة الاحصائية للكشف عن

الأداء التفاضلي للفقرة، كما وأشارت النتائج أيضاً أنه عند استخدام أحجام عينات (500,250) وعند النسبة والتناسب (10:1) فإن الخطأ من النوع سيتضخم، مما أثر على القوة الإحصائية في الكشف عن الكشف عن الأداء التفاضلي للفقرة.

كما أجرى ثيرمان (Thurman, 2009) دراسة لأثر تمييز الفقرة وصعوبتها في الكشف عن الأداء التفاضلي للفقرة في الفقرات متعددة الاجابة. فقد استخدمت طريقة مونت كارلو (monte carlo) لتوليد البيانات. وقد تم إجراء مقارنة للخطأ من النوع الاول (أي أن المؤشرات الإحصائية المستخدمة في هذه الدراسة للكشف عن الأداء التفاضلي بينت أن الفقرة ليست ذات أداء تفاضلي مع أنها ذات أداء تفاضلي)، باستخدام طريقة "مانتل هانزل" (MH)، وطريقة الانحدار اللوجستي، حيث كان التباين في معالم تمييز الفقرة وفي أنماط الأداء التفاضلي للفقرة (منتظم - غير منتظم). وقد أشارت النتائج بوجود علاقة ارتباطية بين ارتفاع مستوى التمييز للفقرة وبين زيادة مستوى الكشف عن الأداء التفاضلي للفقرة، كما وأشارت النتائج أيضاً أن موقع معلمة الصعوبة كان له أثر مباشر في الكشف عن الأداء التفاضلي للفقرة، وأشارت النتائج أيضاً أن مستوى الأداء التفاضلي للفقرة يؤثر في الكشف عن الأداء التفاضلي للفقرة.

التعقيب على الدراسات السابقة:

1. بالنسبة للدراسات التي تناولت الطرق المختلفة للكشف عن الأداء التفاضلي للفقرة تبين أن هنالك تناقض في نتائج الدراسات السابقة فلم يكن هناك طريقة فعالة للكشف عن الأداء التفاضلي، والدليل على ذلك هو عدم الإتساق في نتائج هذه الدراسات؛ ويرجع سبب ذلك إلى أن الأداء التفاضلي للفقرة يعتمد على عدة عوامل منها: حجم العينة، نوع الأداء التفاضلي، طول الاختبار، ومصدر الأداء التفاضلي، ومستوى

تميز الفقرة وموقع معلمة الصعوبة للفقرة، مثل دراسة إنجلهارد وغامر (Engelhard & Gamer, 1997)، ووالستد وروبسون (Walsted & Robson, 1997) والحمداني (2007)، وعبد العزيز (2006)، وقيليان (Gillian, 1989).

2. ركزت الدراسات الأجنبية السابقة التي بحثت في أثر الفقرات ذات الأداء التفاضلي في الاختبار على ما يلي:

أ. أثر الفقرات ذات الأداء التفاضلي على معادلة الاختبار، مثل دراسة زهانغ وماثيوز ودورانز (Zhang & Matthews & Dorans, 2003).

ب. أثر الفقرات التي تظهر أداء تفاضلي على إحصائيات النظرية الكلاسيكية للاختبار (معامل الثبات، الخطأ المعياري في القياس، معامل التصحيح لأثر الإنخفاض)، مثل دراسة زومبو (Zumbo, 2000).

ج. أثر الفقرات ذات الأداء التفاضلي على نسبة الخطأ من النوع الأول، وحجم الأثر باستخدام اختبار (T-test) مثل دراسة زهان وزومبو (Zhan & Zumbo, 2009).

تبين من نتائج هذه الدراسات كما تم عرضها سابقاً؛ أن بعض هذه الإحصائيات قد تأثرت بوجود الفقرات ذات الأداء التفاضلي، والبعض الآخر لم يتأثر بوجود الفقرات ذات الأداء التفاضلي، كما ويلاحظ من هذه الدراسات أنها ركزت بشكل أساسي من حيث أثر الفقرات ذات الأداء التفاضلي على النظرية الكلاسيكية، ولم يتم إيجاد أي دراسة بحثت في أثر الفقرات ذات الأداء التفاضلي على معالم فقرات الاختبار وقدرات الافراد وفقاً لنظرية الاستجابة للفقرة.

3. فيما يخص الدراسات المحلية فقد ركزت هذه الدراسات على المقارنة بين الطرق المختلفة للكشف عن الأداء التفاضلي في ضوء ظروف مختلفة، إلا أن هذه الدراسات - في حدود علم الباحث - لم تتناول أثر

الفقرات ذات الأداء التفاضلي على معالم فقرات الاختبار وقدرات الأفراد في ضوء نظرية الاستجابة للفقرة.

وبعد اطلاع الباحث على الدراسات السابقة لم يتم إيجاد أي دراسة سابقة قامت بشكل مباشر بالبحث في أثر الفقرات ذات الأداء التفاضلي في: مطابقة الأفراد والفقرات للنموذج، وافتراسات النموذج (احادية البعد والاستقلال الموضوعي)، وتقدير معالم القدرة للأفراد، وتقدير معالم الفقرات (الصعوبة، التمييز، التخمين)، وكمية المعلومات التي يقدمها الاختبار. وبناءً على ذلك؛ فقد أضافت هذه الدراسة بعداً جديداً يتمثل بدراسة أثر الفقرات ذات الأداء التفاضلي على احصائيات نظرية الاستجابة للفقرة التي ذكرت سابقاً باستخدام ظروف مختلفة، وهي: نسبة الفقرات ذات الأداء التفاضلي في الاختبار، وهي (10%, 20%, 40%)، ومستوى الأداء التفاضلي للفقرة (متدنٍ بمقدار 0.5 وعالٍ بمقدار 1.5)، حيث تعتبر هذه الظروف من أكثر الظروف التي استخدمت في نتائج الدراسات السابقة، لذلك تعتبر هذه الدراسة إستكمالاً للدراسات السابقة التي لم تتناول بشكل مباشر أثر الفقرات ذات الأداء التفاضلي على معالم فقرات الاختبار وقدرات الأفراد.

الفصل الثالث

الطريقة والإجراءات

يتناول هذا الفصل وصفاً للإجراءات المستخدمة للإجابة عن أسئلة الدراسة من خلال الإعتماد على بيانات مولدة وكيفية توليد البيانات للاستجابة على الفقرات. ويتضمن تصميم توليد البيانات ووصفاً لكل من مستوى الفقرات ذات الأداء التفاضلي، ونسبة الفقرات ذات الأداء التفاضلي في الاختبار. كما تناول وصفاً لكيفية توليد البيانات وإجراءات تقدير المعالم للفقرات وللأفراد. ووصفاً للمعالجات الإحصائية المستخدمة للإجابة عن أسئلة الدراسة حسب الإجراءات التالية:

أولاً: الظروف التي تستند إليها إجراءات توليد البيانات.

1. من حيث تصميم البيانات المستخدم في هذه الدراسة، فقد تكون من (7) مجموعات، وهي مجموعة مرجعية واحدة، وهذه المجموعة لا يوجد بها أي فقرة ذات أداء تفاضلي، وقد كانت هذه المجموعة تمثل بيانات القاعدة الأساسية للمقارنة مع باقي المجموعات، والمجموعات المتبقية جميعها مجموعات مستهدفة تحتوي على فقرات ذات أداء تفاضلي بظروف مختلفة.

2. من حيث عدد الفقرات الكلية في الاختبار، بلغ عددها (50) فقرة في كل مجموعة من المجموعات، وكانت الفقرات التي لا تظهر أداءً تفاضلياً مكررة كما هي في جميع المجموعات، حيث كان الاختلاف بين المجموعات في الفقرات التي تظهر أداءً تفاضلياً، ففي كل مجموعة من المجموعات ظرفية خاصة تختلف بها عن المجموعات الأخرى.

3. عدد الأفراد في كل مجموعة من المجموعات. حيث بلغ عدد الأفراد في كل مجموعة (510) أفراد، وكان التصميم المستخدم في هذه الدراسة يشتمل على (7) مجموعات. وبناءً على ذلك؛ فقد بلغ عدد الأفراد الكلي (3570) مفحوصاً.

4. نسبة الفقرات ذات الأداء التفاضلي الموجودة في الاختبار. لقد تم اختبار ثلاث حالات للاختبار حسب

نسبة الفقرات ذات الأداء التفاضلي في الاختبار، وهي (10%, 20%, 40%).

5. مستوى الأداء التفاضلي الفقرة. أُعتمد في هذه الدراسة حالتين لمستوى الأداء التفاضلي للفقرة، وهي

(مستوى متدنٍ بمقدار 0.5، ومستوى عالٍ بمقدار 1.5)، وذلك حسب الإحصاءات التي قدمها زيكي

(2007).

6. نوع الأداء التفاضلي. اقتصرَت هذه الدراسة على الأداء التفاضلي المنتظم.

إجراءات توليد البيانات

استخدم برنامج Wingen لتوليد معالم (50) فقرة في ظل ظروف مختلفة، وهي مستوى الأداء

التفاضلي (0.5, 1.5)، ونسبة الفقرات ذات الأداء التفاضلي الموجودة في الاختبار، وهي

(10%, 20%, 40%)، كما وتم توليد البيانات متمثلة باستجابات المفحوصين باستخدام النموذج اللوجيستي

ثلاثي المعالم، وركزت الدراسة الحالية على الفقرات ثنائية التدرج، وكانت عملية توليد هذه الاستجابات في

هذه المرحلة مكونة من خطوتين رئيسيتين تم إعتمادها بشكل متكرر من أجل الإجابة عن أسئلة الدراسة.

الخطوة الأولى

في هذه الخطوة تم توليد البيانات الخاصة بالمجموعة المرجعية، حيث تم توليد قيم القدرة للمفحوصين

من خلال استخدام التوزيع المنتظم (uniform)، وكانت أعلى قدرة (maximam) هي (+3)، وأدنى قدرة

(minimam) هي (-3)، وذلك باستخدام النموذج اللوجيستي ثلاثي المعلمة. وتم توليد معلمة الصعوبة

باستخدام التوزيع المنتظم (uniform)، وكانت أدنى قيمة هي (-1.5)، وأعلى قيمة هي (+1.5). كما تم

توليد معلمة التمييز باستخدام التوزيع المنتظم، وكانت أدنى قيمة (0.2) وأعلى قيمة (1.2). واستخدم أيضاً

التوزيع المنتظم في توليد معلمة التخمين حيث كانت أدنى قيمة (0.2)، وأعلى قيمة (0.25). ويبين الجدول

(1) معالم الفقرات التي تم توليدها في المجموعة المرجعية.

الجدول (1): قيم معالم الفقرات ثنائية التدرج من حيث (الصعوبة والتمييز والتخمين) لدى المجموعة المرجعية تبعاً للنموذج اللوجستي ثلاثي المعلمة.

رقم الفقرة	معالم الفقرة		
	التمييز A	الصعوبة B	التخمين C
1	0.212	-0.898	0.222
2	0.224	-0.693	0.25
3	0.265	-1.379	0.215
4	0.294	0.535	0.202
5	0.316	-1.114	0.269
6	0.319	0.299	0.221
7	0.324	0.088	0.265
8	0.331	0.811	0.278
9	0.332	0.939	0.207
10	0.359	0.693	0.262
11	0.405	0.361	0.253
12	0.447	-1.282	0.246
13	0.449	0.66	0.218
14	0.468	0.876	0.286
15	0.496	-0.231	0.226
16	0.508	-0.204	0.22
17	0.528	-1.053	0.271
18	0.531	-0.806	0.264
19	0.562	-0.208	0.229
20	0.57	-0.16	0.249
21	0.592	-0.299	0.24
22	0.603	-1.467	0.295
23	0.613	-0.76	0.276
24	0.614	1.172	0.253
25	0.642	-0.635	0.242
26	0.649	0.334	0.298
27	0.679	-1.456	0.252
28	0.712	-1.358	0.269
29	0.715	-1.373	0.242
30	0.787	-0.711	0.207
31	0.79	0.019	0.25
32	0.839	-0.463	0.27
33	0.896	-0.675	0.27
34	0.9	0.702	0.288
35	0.902	-0.332	0.238
36	0.903	-0.218	0.282
37	0.914	-1.315	0.23
38	0.95	-1.074	0.265
39	0.99	0.166	0.237
40	0.997	0.093	0.228
41	1.001	0.503	0.295
42	1.046	0.778	0.285
43	1.047	1.232	0.238
44	1.106	-1.368	0.289
45	1.123	-0.741	0.289
46	1.15	-1.13	0.289
47	1.171	-0.566	0.287
48	1.179	-1.396	0.266
49	1.195	0.381	0.205
50	1.196	-0.982	0.265

الخطوة الثانية

في الخطوة الثانية، تم توليد بيانات المجموعات المستهدفة باستخدام نفس الإجراءات في الخطوة الأولى باستثناء إجراء تغييرات على قيم معالم الفقرات في المجموعات المستهدفة من أجل أن تعكس أداءً تفاضلياً على بعض الفقرات المختارة، وحسب كل مجموعة من مجموعات البيانات التي تم توليدها. ويبين الجدول (2) أرقام الفقرات التي تم اختيارها من المجموعة المرجعية الجدول (1) لكي تظهر أداءً تفاضلياً في المجموعات المستهدفة، فكانت على النحو التالي:

الجدول(2):أرقام الفقرات ذات الأداء التفاضلي وفقاً لنسبتها في الاختبار (10%,20%,40%).

النموذج الأول الذي يشتمل على 10% من الفقرات ذات الأداء التفاضلي بلاضافة إلى أرقام هذه الفقرات	النموذج الثاني الذي يشتمل على 20% من الفقرات ذات الأداء التفاضلي بلاضافة إلى أرقام هذه الفقرات	النموذج الثالث الذي يشتمل على 40% من الفقرات ذات الأداء التفاضلي بلاضافة إلى أرقام هذه الفقرات
1	1	1
3		
6	6	
8	8	
14		
15		
18	18	
21	21	
23		
26		
28		
29	29	29
32		
34		
35	35	35
39	39	39
40		
42	42	42
44	44	
50		

كيفية اختيار الفقرات الموجودة في الجدول (2) لكي تظهر أداءً تفضلياً .

لقد تم اختيار الفقرات لتظهر أداءً تفضلياً من خلال ترتيب معاملات الصعوبة للفقرات في المجموعة المرجعية تصاعدياً من الأدنى إلى الأعلى، ومن ثم تم تقسيمها إلى خمسة مستويات، واختير من كل مستوى الفقرات التي تتوسط الترتيب من حيث الصعوبة لكي تظهر أداءً تفضلياً، كما هو مبين في الجدول (2).

وبعد إختيار الفقرات لكي تظهر أداءً تفضلياً تم توليد الأداء التفاضلي لهذه الفقرات من خلال الزيادة في قيمة معلمة الصعوبة (b) في المجموعة المستهدفة. ونتيجةً لذلك فإن هذه الفقرة تكون أصعب للمجموعة المستهدفة عبر جميع مستويات القدرة ولصالح المجموعة المرجعية. وفي هذا الموقف يكون هنالك عدة فقرات ذات أداء تفضلي ونسب هذه الفقرات في الاختبار هي (10%, 20%, 40%). ويكون هنالك أيضاً مستويات مختلفة من الأداء التفاضلي للفقرة وهي (متدنٍ بمقدار 0.5، وعالٍ بمقدار 1.5). وفي هذه المواقف وحين يتوفر عدد معين من الفقرات ذات الأداء تفضلي في كل مجموعة، فإن هذه الفقرات تمتلك نفس المستوى من الأداء التفاضلي، وهذا يعنى من حيث سهولة تفسير البيانات فلم يتم البحث في أثر وجود مستويات مختلفة للأداء التفاضلي للفقرة عند دراسة كل مجموعة منفردة تبعاً لنسبة الفقرات ذات الأداء التفاضلي.

وبناءً على ذلك؛ فإن التصميم المستخدم في هذه الدراسة قد تكون من (6) مجموعات مستهدفة كل مجموعة لها خصائص مختلفة، كما هو مبين في الجدول (3)، بالإضافة إلى دراسة مجموعة واحدة تخلو من أي فقرات ذات أداء تفضلي، وتمثل المجموعة المرجعية.

الجدول (3).نسبة الفقرات ذات الأداء التفاضلي، ومستوى الأداء التفاضلي للفقرة لدى المجموعات المستهدفة.

نسبة الفقرات ذات الأداء التفاضلي	مستوى الأداء التفاضلي للفقرة	
	مستوى عالٍ بمقدار 1.5	مستوى متدنٍ بمقدار 0.5
10%	C2	C1
20%	C4	C3
40%	C6	C5

يمثل الرمز C_i (Cisimulation conditions) مجموعات توليد البيانات في الدراسة أو تصميم

الدراسة، الذي يتكون من (7) مجموعات توليد للبيانات، وهي:

❖ ست مجموعات مستهدفة تشتمل على فقرات ذات أداء التفاضلي تتمثل في (3 حالات لنسب الفقرات

ذات الأداء التفاضلي * حالتين لمستوى الأداء التفاضلي للفقرة)

❖ مجموعة واحدة لا تشتمل على وجود أي فقرة ذات أداء تفاضلي، والتي تمثل القاعدة الأساسية للمقارنة

(المجموعة المرجعية).

فعلى سبيل المثال تشير المجموعات التالية إلى:

C2: تشير إلى المجموعة التي تشتمل على (10%) من الفقرات ذات الأداء تفاضلي، ومستوى

الأداء التفاضلي (1.5).

C6: تشير إلى المجموعة التي تشتمل على (40%) من الفقرات ذات الأداء تفاضلي، ومستوى

الأداء التفاضلي (1.5).

وللإجابة عن أسئلة الدراسة

تم اعتماد الظروف التي ذكرت سابقاً، وهي ثلاثة ظروف لنسب الفقرات ذات الأداء التفاضلي في الاختبار، وهي (10%, 20%, 40%)، ومستويين للأداء التفاضلي للفقرة، وهما (0.5, 1.5). وقد تمت مقارنة جميع هذه المجموعات، والتي تمثل المجموعات المستهدفة مع المجموعة المرجعية، وذلك للإجابة عن جميع أسئلة الدراسة.

وللإجابة عن السؤال الأول فقد تم اتباع الخطوات التالية:

لقد تمت مقارنة الفقرات المطابقة والأفراد المطابقين للنموذج اللوجستي ثلاثي المعلمة في المجموعات المستهدفة بالفقرات المطابقة والأفراد المطابقين للنموذج اللوجستي ثلاثي المعلمة في المجموعة المرجعية في ظل جميع الظروف التي ذكرت سابقاً، وذلك من خلال استخدام اختبار كاي تربيع (χ^2).

وللإجابة عن السؤال الثاني

لقد تمت مقارنة معالم الفقرات المقدرة (الصعوبة والتمييز والتخمين) في المجموعات المستهدفة وفقاً (لنسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة) مع معالم الفقرات المقدرة في المجموعة المرجعية من حيث:

1. مقارنة الأوساط الحسابية والانحرافات المعيارية لقيم معالم الفقرات (الصعوبة والتمييز والتخمين) في المجموعات المستهدفة مع الأوساط الحسابية والانحرافات المعيارية لقيم معالم الفقرات في المجموعة المرجعية.

2. إجراء تحليل التباين الثنائي لكل معلمة من معالم الفقرات (الصعوبة والتمييز والتخمين) وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة والتفاعل بينهما.

3. التأكد من خاصية اللاتغير لمنحنيات خصائص الفقرة، وذلك بمقارنة منحنيات خصائص الفقرات التي لا

تظهر أداءً تفاضلياً في المجموعات المستهدفة مع منحنيات خصائص الفقرات في المجموعة المرجعية.

وللإجابة عن السؤال الثالث فقد تم اتباع الخطوات التالية:

1. مقارنة الأوساط الحسابية والانحرافات المعيارية لقدرات الأفراد في المجموعات المستهدفة مع الأوساط

الحسابية والانحرافات المعيارية لقدرات الأفراد في المجموعة المرجعية.

2. إجراء تحليل التباين الثنائي للمتوسطات الحسابية لقدرات الأفراد وفقاً لنسبة الفقرات ذات الأداء التفاضلي

ومستوى الأداء التفاضلي للفقرة والتفاعل بينهما.

3. مقارنة الأوساط الحسابية والانحرافات المعيارية للعلامات الملاحظة للأفراد في المجموعات المستهدفة مع

الأوساط الحسابية والانحرافات المعيارية للعلامات الملاحظة للأفراد في المجموعة المرجعية.

4. إجراء تحليل التباين الثنائي للمتوسطات الحسابية للعلامات الملاحظة للأفراد وفقاً لنسبة الفقرات ذات

الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة والتفاعل بينهما.

5. استخدام اختبار المقارنات البعدية لتحديد أي المجموعات التي أظهرت دلالات إحصائية.

وللإجابة عن السؤال الرابع فقد تم اتباع الخطوات التالية:

1. لقد تم مقارنة معاملات الثبات الأمبريقي للاختبار لدى الأفراد في كل مجموعة من المجموعات

المستهدفة مع المجموعة المرجعية، وتم استخدام الإحصائي M للعينات المستقلة للكشف عن دلالة الفرق

بين قيم معاملي الثبات الأمبريقي.

2. مقارنة المتوسطات الحسابية والانحرافات المعيارية لكمية المعلومات التي يقدمها الاختبار في كل

مجموعة من المجموعات المستهدفة مع المجموعة المرجعية.

3. مقارنة منحنيات دالة معلومات الاختبار مع الخطأ المعياري له في كل مجموعة من المجموعات

المستهدفة مع المجموعة المرجعية .

البرامج الإحصائية المستخدمة في الدراسة الحالية.

• استخدم في هذه الدراسة برنامج spss لإجراء التحليلات التالية:

1. حساب الإحصائيات الوصفية لكل مجموعة من المجموعات.

2. إجراء تحليل التباين الثنائي.

3. استخدام اختبارات المقارنات البعدية.

4. استخدام اختبار كآي تربيع للاستقلالية.

• البرنامج الإحصائي R-fit-3pv-2 (lopez, 2008).

استخدم برنامج (R-Fit-3p v-2) في التأكد من الفقرات المطابقة والأفراد المطابقين للنموذج

اللوجستي ثلاثي المعلمة، حيث تقوم هذه البرمجية في حساب مؤشرات المطابقة على مستوى الفرد والفقرة

في النموذج الأحادي والثنائي والثلاثي المعلمة بنوعها اللوجيستي والطبيعي، ومن المؤشرات الإحصائية

التي تقوم بحسابها (LZ) الذي يستخدم مع النموذج الثلاثي، وهو مثال على الأسلوب المبني على طريقة

الارجحية العظمى.

• البرنامج الإحصائي Bilog-mgv3

استخدم برنامج (Bilog-mg) لإجراء التحليلات التالية:

1. تقدير معالم الفقرات (الصعوبة والتمييز والتخمين) والخطأ المعياري لتقدير هذه المعالم في كل

مجموعة من المجموعات المستخدمة في هذه الدراسة.

2. حساب معالم القدرة والخطأ المعياري لتقدير هذه المعالم في كل مجموعة.

3. حساب كمية المعلومات للفقرات وللاختبار في كل مجموعة.

الفصل الرابع

عرض النتائج ومناقشتها

هدفت هذه الدراسة إلى بيان أثر الفقرات ذات الأداء التفاضلي على كل من: مطابقة الفقرات والأفراد للنموذج اللوجستي ثلاثي المعلمة، وافتراضات النموذج (احادية البعد والاستقلال الموضوعي)، وتقدير معالم فقرات الاختبار وقدرات الأفراد، وكمية معلومات الاختبار، من خلال استخدام بيانات مولدة لتكوين (6) مجموعات مستهدفة تختلف فيما بينها بنسبة الفقرات ذات الأداء التفاضلي في الاختبار، ومستوى الأداء التفاضلي للفقرة، بالإضافة إلى المجموعة المرجعية. ويتناول هذا الفصل عرضاً مفصلاً للنتائج ومناقشتها التي توصلت إليها الدراسة.

النتائج المتعلقة بالسؤال الأول

للإجابة عن السؤال الأول الذي نصّ على: "ما مدى مطابقة الأفراد ومطابقة الفقرات للنموذج اللوجستي ثلاثي المعلمة تبعاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة؟" تم القيام بالخطوات الآتية للتحقق من إجراءات المطابقة وفقاً للنموذج اللوجستي ثلاثي المعلمة:

أ. مطابقة الأفراد لبيانات الاختبار في كل مجموعة وفقاً لنسبة الفقرات ذات الأداء التفاضلي ولمستوى الأداء التفاضلي للفقرة، في ضوء النموذج اللوجستي ثلاثي المعلمة:

استخدم برنامج (Bilog-MG3) للكشف عن مطابقة الأفراد تبعاً لمتغير نسبة الفقرات ذات الأداء التفاضلي في الاختبار، وهي (10%, 20%, 40%)، ومستوى الأداء التفاضلي (0.5, 1.5). حيث تم حذف الأفراد غير المطابقين في كل مجموعة، ثم تم التحقق من دلالة الفرق بين نسب الأفراد المطابقين للنموذج الثلاثي المعلمة في نظرية الاستجابة للفقرة ما بين مجموعات الدراسة، كما في الجدول (4).

الجدول (4): أعداد وأرقام الأفراد غير المطابقين في كل مجموعة وفقاً لنسبة الفقرات ذات الأداء التفاضلي ولمستوى الأداء التفاضلي للفقرة، ونتائج الفرق بين نسب الأفراد المطابقين عبر جميع المجموعات.

حالة الفقرة	تكرار الأفراد غير المطابقين	أرقام الأفراد غير المطابقين	التكرار المتوقع	الباقى	كاي ² الفرعية	درجة الحرية	الدلالة الإحصائية
المرجعية	1	379	1.57	-0.57	0.21	1	0.649
المستهدفة (النسبة 10% ، المستوى 0.5)	0		1.57	-1.57	1.57	1	0.210
المستهدفة (النسبة 10% ، المستوى 1.5)	2	299, 387	1.57	0.43	0.12	1	0.732
المستهدفة (النسبة 20% ، المستوى 0.5)	0		1.57	-1.57	1.57	1	0.210
المستهدفة (النسبة 20% ، المستوى 1.5)	0		1.57	-1.57	1.57	1	0.210
المستهدفة (النسبة 40% ، المستوى 0.5)	4	2, 140, 251, 490	1.57	2.43	3.75	1	0.053
المستهدفة (النسبة 40% ، المستوى 1.5)	4	275, 298, 363, 430	1.57	2.43	3.75	1	0.053
الكل	11		11	0	12.55	6	0.051

أظهرت نتائج التحليل المبين في الجدول (4) أعداد وأرقام الأفراد غير المطابقين في كل مجموعة حسب نسبة الفقرات ذات الأداء التفاضلي، ومستوى الأداء التفاضلي للفقرة، حيث تبين من خلال الجدول أن أعداد الأفراد غير المطابقين لا يتجاوز (4) أفراد من أصل (510) أفراد في كل مجموعة. كما يُلاحظ بأنه لم يكن هنالك تقاطعاً بأرقام الأفراد غير المطابقين عبر جميع المجموعات، فقد تم حذف الأفراد غير المطابقين في كل مجموعة، وذلك بسبب عدم مطابقتهم للنموذج ثلاثي المعلمة لتعذر حساب الخطأ المعياري في تقدير القدرة لهم عند مستوى قدراتي ما.

كما يُلاحظ من خلال الجدول (4)، أنه لا توجد فروق ذات دلالة إحصائية بين نسب الأفراد المطابقين للنموذج الثلاثي المعلمة في نظرية الاستجابة للفقرة بين جميع المجموعات تعزى لنسبة الفقرات ذات الأداء التفاضلي ولمستوى الأداء التفاضلي للفقرة، حيث بلغت احتمالية الدلالة الإحصائية (0.051).
ب. مطابقة الفقرات في كل مجموعة وفقاً لنسبة الفقرات ذات الأداء التفاضلي ولمستوى الأداء التفاضلي للفقرة، في ضوء النموذج ثلاثي المعلمة.

تم التحقق من مطابقة فقرات الاختبار لنموذج الاستجابة للفقرة ثلاثي المعلمة للبيانات المولدة في كل مجموعة من المجموعات تبعاً لنسبة الفقرات ذات الأداء التفاضلي ولمستوى الأداء التفاضلي للفقرة، حيث تم

حذف الفقرات غير المطابقة من بيانات الاختبار في كل مجموعة، والجدول (5) يبين إحصائيات الفقرات غير المطابقة.

الجدول (5): إحصائيات مطابقة الفقرات للمجموعات التي ظهر لديها فقرات غير مطابقة وفقاً لنسبة الفقرات ذات الأداء التفاضلي ولمستوى الأداء التفاضلي للفقرة.

رقم الفقرة غير مطابقة	معلمة التمييز	الخطأ المعياري لمعلمة التمييز	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة	معلمة التخمين	الخطأ المعياري لمعلمة التخمين	χ^2	درجة الحرية	احتمالية المطابقة	حالة الفقرة في المجموعة:
22	1.145	0.220	-0.276	0.318	0.365	0.085	36.1	8	0.000	المستهدفة (النسبة 10%، المستوى 0.5)
1	1.476	0.459	1.126	0.218	0.487	0.045	26.5	9	0.002	المستهدفة (النسبة 20%، المستوى 0.5)
37	1.890	0.297	-0.516	0.184	0.313	0.068	36.5	8	0.000	المستهدفة (النسبة 40%، المستوى 0.5)
16	1.128	0.285	0.440	0.298	0.382	0.074	22.0	9	0.009	المستهدفة (النسبة 40%، المستوى 1.5)
50	2.124	0.346	0.282	0.119	0.242	0.045	24.3	8	0.002	

يُلاحظ من نتائج مطابقة الفقرات في جدول (5)، أن هنالك فقرة واحدة لم تطابق في المجموعة المستهدفة التي تتمثل في أن (10% من فقراتها ذات أداء تفاضلي، ومستوى الأداء التفاضلي لهذه الفقرات 0.5)، وتم حذفها من المجموعة وكان رقمها (22). كما تم حذف الفقرة رقم (1) في المجموعة المستهدفة التي تتمثل في أن (20% من فقراتها ذات أداء تفاضلي، ومستوى الأداء التفاضلي لهذه الفقرات 0.5) بسبب عدم مطابقتها، وحذفت الفقرة رقم (37) في المجموعة المستهدفة التي تتمثل في أن (40% من فقراتها ذات أداء تفاضلي، ومستوى الأداء التفاضلي لهذه الفقرات 0.5)، وتم حذف فقرتين رقمهما (16,50) في المجموعة المستهدفة التي تتمثل في أن (40% من فقراتها ذات أداء تفاضلي، ومستوى الأداء التفاضلي لهذه الفقرات 0.5)، حيث كانت احتمالية المطابقة لجميع هذه الفقرات أقل من (0.01)، وهذا مؤشر على عدم مطابقة هذه الفقرات. وبالنسبة لباقي المجموعات فقد كانت جميع فقراتها مطابقة للنموذج.

كما يُلاحظ من الجدول (5) أن أعداد الفقرات غير المطابقة تراوحت عبر جميع المجموعات من (2-0) فقرة من أصل (50) فقرة في كل مجموعة، ويتبين من خلال ذلك أنه لا يوجد فروق في مطابقة الفقرات بين

المجموعات تُعزى إلى نسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة. وتدل هذه النتائج المتعلقة بمطابقة الفقرات والأفراد للنموذج اللوجستي ثلاثي المعلمة، على أن نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة، لم يكن لهما أثر دال إحصائياً على مطابقة الفقرات والأفراد. كما تؤكد هذه النتيجة قوة الإجراءات لمطابقة الفقرات والأفراد التي تقوم عليها نظرية الاستجابة للفقرة، ولأن البيانات أصلاً تم توليدها بما يتطابق مع النموذج ثلاثي المعلمة.

النتائج المتعلقة بالسؤال الثاني

للإجابة عن السؤال الثاني الذي نصَّ على: "ما مدى تحقق افتراضات النموذج (أحادية البعد والاستقلال الموضعي) في الاختبار تبعاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة؟"؛ تم القيام بالخطوات الآتية للتحقق من افتراضات نظرية الاستجابة للفقرة في كل مجموعة من المجموعات المستهدفة بالإضافة إلى المجموعة المرجعية وفقاً للنموذج اللوجستي ثلاثي المعلمة:

1. التحقق من افتراض أحادية البعد (Unidimensionality) لبيانات الإختبار في كل مجموعة من

المجموعات المستهدفة والمجموعة المرجعية وفقاً لنسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي.

يعد افتراض أحادية البعد من أهم الافتراضات في نظرية الاستجابة للفقرة، وللتحقق من هذا الافتراض تم تحليل

بيانات الإختبار في كل مجموعة، باستخدام برنامج (Normal Ogive Harmonic Analysis) NOHARM

(Robust Method) ، ويبين الجدول (6) نتائج التحليل العاملي لكل مجموعة.

الجدول (6): نتائج التحليل العاملي لفقرات الإختبار النهائية في كل مجموعة وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الإختبار

ولمستوى الأداء التفاضلي.

حالة الفقرات	مجموع مربعات البواقي (دون القطر الرئيسي)	الجذر التربيعي لمتوسط مجموع مربعات البواقي (دون القطر الرئيسي)	مؤشر Tanaka لحسن المطابقة	عدد الأفراد المطابقين
المرجعية	0.09285	0.00871	0.18173	509
المستهدفة (النسبة 10%، المستوى 0.5)	0.08520	0.00851	0.18155	510
المستهدفة (النسبة 10%، المستوى 1.5)	0.09422	0.00877	0.18191	508
المستهدفة (النسبة 20%، المستوى 0.5)	0.08697	0.00860	0.18155	510
المستهدفة (النسبة 20%، المستوى 1.5)	0.09066	0.00860	0.18155	510
المستهدفة (النسبة 40%، المستوى 0.5)	0.09583	0.00903	0.18227	506
المستهدفة (النسبة 40%، المستوى 1.5)	0.08637	0.00875	0.18227	506

أظهر التحليل في الجدول (6) بيانات عينة التطبيق النهائي المطابقة للنموذج ثلاثي المعلمة والمتعلقة بإستجابات الأفراد في كل مجموعة من المجموعات المستهدفة والمجموعة المرجعية، وتم التحقق من إفتراض أحادية البعد في كل مجموعة من خلال مؤشرين هما: مؤشر تاناكا (TANAKA)، فقد كانت أقل قيمة لهذا المؤشر (0.977)، وكانت لدى المجموعة المرجعية، وأعلى قيمة لهذا المؤشر (0.9885) وكانت لدى المجموعة المستهدفة التي تتمثل في أن (20% من فقراتها ذات أداء تفاضلي ومستوى الاداء التفاضلي لهذه الفقرات 1.5)، وهما قيمتان تشيران إلى تحقق إفتراض أحادية البعد وفقاً لـ (Jaspar, 2010)، وحيث كانت جميع القيم أكبر من (0.94)، والمؤشر الآخر؛ هو مؤشر الجذر التربيعي لمتوسط مربعات البواقي (RMSR)، وقد بلغت أقل قيمة محسوبة هي (0.00850) وكانت لدى المجموعة المرجعية وهي أقل من القيمة الحرجة لها البالغة قيمتها (0.18190) التي تحسب من خلال المعادلة $(4.1/\sqrt{509})$ ، وكذلك بلغت أعلى قيمة محسوبة لـ (RMSR) هي (0.0090) لدى المجموعة المستهدفة التي تتمثل في أن (40% من فقراتها ذات أداء تفاضلي ومستوى الاداء التفاضلي لهذه الفقرات 0.5)، وهي أقل من القيمة الحرجة له البالغة قيمتها (0.1817) التي تحسب من خلال المعادلة $(4.1/\sqrt{509})$ ، وعند مقارنة القيمتان الناتجتان من خلال التحليل مع أعلى قيمة لهذا المؤشر تبين أن القيمتان الناتجتان تقتربان من الصفر وهذا مؤشر على تحقق إفتراض أحادية البعد (Jasper, 2010). ويلاحظ من

خلال الجدول أن أعلى فرق بين جميع الحالات لمؤشر تاناكا كانت أقل من (0.01) بمعنى لم يكن هنالك فروق ظاهرية لمؤشر أحادية البعد بين جميع المجموعات المستهدفة والمجموعة المرجعية عزى لنسبة الفقرات ذات الأداء التفاضلي ولمستوى الأداء التفاضلي للفقرة.

2. التحقق من افتراض الإستقلال الموضوعي (Local Independence) لبيانات الإختبار في كل مجموعة من المجموعات المستهدفة والمجموعة المرجعية وفقاً لنسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة.

لأغراض التحقق من افتراض الإستقلال الموضوعي؛ تم إستخدام برنامج LDID (Local Dependence Indices for Dichotomous Items) على بيانات العينة النهائية المطابقة للنموذج الثلاثي في كل مجموعة من المجموعات المستهدفة والمجموعة المرجعية، وذلك كما هو مبين في الجدول (7).

الجدول (7): الإحصائيات الخاصة بمؤشر ZQ3 للإستقلال الموضوعي في كل مجموعة من المجموعات المستهدفة والمجموعة المرجعية وفقاً لنسبة الفقرات ذات الأداء التفاضلي ولمستوى الأداء التفاضلي للفقرة.

حالة الفقرات	المتوسط الحسابي المشاهد	الانحراف المعياري لمشاهد	المتوسط الحسابي النظري	التباين -Zq3	الانحراف المعياري -Zq3	الانحراف المعياري -Zq3 *2	فترة الثقة		عدد الفقرات المطابقة	عدد الأفراد المطابقين
							الحد الأدنى	الحد الأعلى		
المرجعية	-0.0074147	0.048081816	-0.020408	0.001976	0.044455	0.088911	-0.096326	0.081496	50	509
المستهدفة (النسبة 10% ، المستوى 0.5)	-0.00667119	0.050854257	-0.020833	0.001972	0.044412	0.088823	-0.095494	0.082152	49	510
المستهدفة (النسبة 10% ، المستوى 1.5)	-0.00755962	0.048224791	-0.020408	0.001980	0.044499	0.088999	-0.096558	0.081439	50	508
المستهدفة (النسبة 20% ، المستوى 0.5)	-0.00723218	0.05313507	-0.020833	0.001972	0.044412	0.088823	-0.096055	0.081591	49	510
المستهدفة (النسبة 20% ، المستوى 1.5)	-0.00943957	0.046295728	-0.020408	0.001972	0.044412	0.088823	-0.098263	0.079384	50	510
المستهدفة (النسبة 40% ، المستوى 0.5)	-0.00787777	0.04904328	-0.020833	0.001988	0.044588	0.089176	-0.097053	0.081298	49	506
المستهدفة (النسبة 40% ، المستوى 1.5)	-0.01054329	0.061396614	-0.021277	0.001988	0.044588	0.089176	-0.099719	0.078632	48	506

تبين من التحليل في الجدول (7) عبر جميع المجموعات أن أعلى قيمة للوسط الحسابي لمؤشر ZQ3 كان لدى المجموعة المستهدفة التي تتمثل في أن (40% من فقراتها ذات أداء تفاضلي ومستوى الاداء التفاضلي لهذه الفقرات 1.5)، حيث بلغت قيمته (-0.0105) والتي تقع بين الحد النظري الأدنى لفترة الثقة لقيم ZQ3 البالغة قيمة (-0.099719) وبين الحد النظري الأعلى لفترة الثقة لقيم (ZQ3) البالغة قيمة (0.078632)؛ كما تبين من التحليل أن أدنى قيمة للوسط الحسابي لمؤشر (ZQ3) كان لدى المجموعة المستهدفة التي تتمثل في أن (10% من فقراتها ذات

أداء تفاضلي ومستوى الأداء التفاضلي لهذه الفقرات (0.5)، وبلغت قيمته (-0.00667119)، والتي تقع بين الحد النظري الأدنى لفترة الثقة لقيم ZQ3 البالغة قيمته (-0.09549)، وبين الحد النظري الأعلى لفترة الثقة لقيم ZQ3 البالغة قيمتها (0.082152)؛ كما تم حساب عدد أزواج الفقرات التي وقعت ضمن فترة الثقة المحققة لشرط الإستقلال الموضعي للأفراد في كل مجموعة من المجموعات المستهدفة والمجموعة المرجعية، والجدول (8) يبين مؤشرات الإستقلال الموضعي وفقاً لنظرية الإستجابة للفقرة.

الجدول (8). مؤشرات الإستقلال الموضعي في كل مجموعة من المجموعات المستهدفة والمجموعة المرجعية وفقاً لنسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة.

الدالة الإحصائية	درجة الحرية	2ا الفرعية	الباقى	النسبة المتوقعة	الكلي		حالة الاستقلال الموضوعي				حالة الفقرة
					النسبة المئوية	عدد الأزواج	النسبة المئوية	عدد الأزواج	النسبة المئوية	عدد الأزواج	
0.862	1	0.030	1.67	92.1	100.0	1225	93.8	1149	6.2	76	المرجعية
0.998	1	0.000	-0.03	92.1	100.0	1176	92.1	1083	7.9	93	(النسبة 10%، المستوى 0.5)
0.870	1	0.027	1.57	92.1	100.0	1225	93.7	1148	6.3	77	(النسبة 10%، المستوى 1.5)
0.873	1	0.025	-1.53	92.1	100.0	1176	90.6	1065	9.4	111	(النسبة 20%، المستوى 0.5)
0.829	1	0.047	2.07	92.1	100.0	1225	94.2	1154	5.8	71	(النسبة 20%، المستوى 1.5)
0.944	1	0.005	0.67	92.1	100.0	1176	92.8	1091	7.2	85	(النسبة 40%، المستوى 0.5)
0.645	1	0.213	-4.43	92.1	100.0	1128	87.7	989	12.3	139	(النسبة 40%، المستوى 1.5)
0.999	6	0.347	0.000	644.9			644.9	7679			الكلي

يُلاحظ من الجدول (8)، أن النتائج الخاصة به قد كانت على النحو الآتي:

عدم وجود إنتهاك لافتراض الاستقلال الموضعي للفقرات في كل المجموعات، فقد كانت أعلى قيمة لعدد الأزواج المرتبطة لدى المجموعة المستهدفة التي تتمثل في أن (0.40 من فقراتها ذات أداء تفاضلي ومستوى الاداء التفاضلي لهذه الفقرات 1.5)، فقد بلغ عدد أزواج الفقرات التي وقعت خارج المدى (139) بنسبة (12.3%)، بينما كان عدد أزواج الفقرات المستقلة التي وقعت ضمن المدى (989) بنسبة (87.7%)، ويتبين من هذا أن عدد أزواج الفقرات التي حققت الإستقلالية أعلى بـ (8) أضعاف تقريباً من عدد أزواج الفقرات التي حققت التبعية الموضعية، وهذا مؤشر قوي على تحقق افتراض الإستقلال الموضعي. وكانت أدنى قيمة لعدد الأزواج المرتبطة لدى المجموعة المستهدفة التي تتمثل في أن (20% من فقراتها ذات أداء تفاضلي ومستوى

الاداء التفاضلي لهذه الفقرات (1.5)، فقد بلغت عدد أزواج الفقرات التي وقعت خارج المدى (71) زوج بنسبة (5.8%)، بينما كان عدد أزواج الفقرات المستقلة التي وقعت ضمن المدى (1154) بنسبة (94.2%)، ويتبين من هذا أن عدد أزواج الفقرات التي حققت الإستقلالية أعلى بـ (16) ضعف تقريباً من عدد أزواج الفقرات التي حققت التبعية الموضوعية، وهذا مؤشر قوي على تحقق إفتراض الإستقلال الموضوعي. كما يُلاحظ من خلال الجدول عدم وجود فروق ذات دلالات إحصائية لمؤشرات الإستقلال الموضوعي بين جميع المجموعات تُعزى لنسبة الفقرات ذات الاداء التفاضلي ولمستوى الأداء التفاضلي للفقرة، حيث كانت قيمة إحصائية الدلالة الإحصائية (0.99). وتدل هذه النتائج على عدم وجود إنتهاك لافتراض الاستقلال الموضوعي وأحادية البعد يُعزى لنسبة الفقرات ذات الاداء التفاضلي ولمستوى الأداء التفاضلي للفقرة، ويعود السبب في ذلك إلى قوة هذه الافتراضات التي تستند اليها نظرية الاستجابة للفقرة.

النتائج المتعلقة بالسؤال الثالث

للإجابة عن السؤال الثالث الذي نصَّ على: "هل توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين متوسطات تقديرات معالم فقرات الاختبار (الصعوبة والتمييز والتخمين) تُعزى لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة أو التفاعل بينهما؟"، فقد تم تقدير معالم فقرات الاختبار (الصعوبة والتمييز والتخمين) والأخطاء المعيارية لها في كل مجموعة من المجموعات المستهدفة والمجموعة المرجعية، علاوةً على حساب قيمة اختبار χ^2 لحسن المطابقة لكل فقرة من فقرات الاختبار، وذلك كما هو موضح في الملاحق (أ،ب،ت،ث،ج،ح،خ).

ولأغراض المقارنة فقد تم إيجاد قيم المتوسطات الحسابية والانحرافات المعيارية لكل معلمة من معالم فقرات الاختبار (التمييز والصعوبة والتخمين)، بالإضافة الى إجراء تحليل التباين الثنائي، وذلك على النحو التالي:

فيما يخص معلمة الصعوبة

تم حساب المتوسطات الحسابية والانحرافات المعيارية لمعلمة الصعوبة لفقرات الاختبار في كل مجموعة حسب نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة، وذلك كما هو مبين في الجدول (9).

الجدول (9) . الوسط الحسابي والانحراف المعياري لمعلمة الصعوبة لفقرات الاختبار في كل مجموعة حسب نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.

نسبة الفقرات ذات الاداء التفاضلي في الاختبار	مستوى الاداء التفاضلي للفقرة	المتوسط الحسابي لمعلمة الصعوبة	الانحراف المعياري لمعلمة الصعوبة
المرجعية	المرجعية	0.273	0.41
الكلية	الكلية	0.273	0.41
10%	0.5	0.344	0.48
	1.5	0.420	0.56
الكلية	الكلية	0.382	0.52
20%	0.5	0.312	0.50
	1.5	0.493	0.58
الكلية	الكلية	0.403	0.54
40%	0.5	0.540	0.49
	1.5	0.672	0.78
الكلية	الكلية	0.606	0.65
الكلية	المرجعية	0.382	0.41
	0.5	0.419	0.50
	1.5	0.477	0.66

يلاحظ من الجدول (9)، وجود فروق ظاهرية بين المتوسطات الحسابية لمعلمة الصعوبة لفقرات الاختبار ناتجة عن اختلاف مجموعات متغيري الدراسة، حيث كان أعلى قيمة للوسط الحسابي والانحراف المعياري لمعلمة الصعوبة لفقرات الاختبار في المجموعة المستهدفة التي تتمثل في أن (40% من فقراتها ذات أداء تفاضلي ومستوى الأداء التفاضلي لهذه الفقرات 1.5)، وبلغت القيم على الترتيب (0.672, 0.78)، وكانت أقل قيمة للوسط الحسابي والانحراف المعياري لمعلمة الصعوبة عند المجموعة المرجعية، وقد بلغت القيم بنفس الترتيب (0.273, 0.41)، بفارق مقداره بنفس الترتيب (0.4, 0.37). كما يُلاحظ من خلال

الجدول بأنه كلما زاد نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرات زادت قيمة الوسط الحسابي والانحراف المعياري لمعلمة الصعوبة، وللتحقق من جوهرية الفروق الظاهرية؛ تم إجراء تحليل التباين الثنائي للمتوسطات الحسابية لمعلمة الصعوبة لفقرات الاختبار وفقاً لمتغيري الدراسة، وذلك كما في الجدول (10).

الجدول (10). نتائج تحليل التباين الثنائي للمتوسطات الحسابية لمعلمة الصعوبة حسب نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرات.

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
نسبة الفقرات ذات الأداء التفاضلي	3.803	2	1.902	6.242	0.002
مستوى الأداء التفاضلي للفقرات	0.268	1	0.268	0.880	0.349
التفاعل	1.400	2	0.700	2.298	0.102
الخطأ	102.965	338	0.305		
الكل	108.603	344			

يتضح من الجدول (10)، وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين المتوسطات الحسابية لمعلمة الصعوبة تُعزى لنسبة الفقرات ذات الأداء التفاضلي في الاختبار. أما مستوى الأداء التفاضلي للفقرات والتفاعل بين المتغيرين فلم يكن ذا دلالة إحصائية. ولكون المتغير (نسبة الفقرات ذات الأداء التفاضلي) متعدد المجموعات؛ فقد تم إجراء اختبار Levene للتحقق من إنتهاك تجانس التباين للمتوسطات الحسابية لمعلمة الصعوبة؛ بهدف تحديد أنسب اختبار مقارنات بعدية للكشف عن جوهرية الفروق بين المتوسطات الحسابية لمعلمة الصعوبة. وقد بلغت قيمة ف المحسوبة الخاصة باختبار Levene (3.788) وهي قيمة دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) عند درجتي حرية (6 للبسط و338 للمقام)، مما يعني وجود إنتهاك لتجانس التباين لمعلمة الصعوبة يُعزى لنسبة الفقرات ذات الأداء التفاضلي في الاختبار؛ بما يفيد ضرورة إجراء اختبار Games-Howell للمقارنات البعدية المتعددة للكشف عن

جوهرية الفروق بين المتوسطات الحسابية لمعلمة الصعوبة وفقاً لنسبة الفقرات ذات الأداء التفاضلي، وذلك كما في الجدول (11).

الجدول (11). نتائج اختبار Games-Howell للكشف عن مواقع الفروق بين المتوسطات الحسابية لمعلمة الصعوبة حسب نسبة الفقرات ذات الأداء التفاضلي في الاختبار.

نسبة الفقرات ذات الأداء التفاضلي		Games-Howell		المرجعية	%10	%20	%40
المرجعية		0.273		0.273	0.382	0.403	0.606
%10		0.109		0.130	0.224	0.203	
%20		0.403		0.267			
%40		0.606					

يتضح من الجدول (11) بأنه لا يوجد إلا فرقاً واحداً دال إحصائياً بين المتوسطات الحسابية لمعلمة الصعوبة وكان بين المجموعة المستهدفة التي تشتمل على 40% من الفقرات ذات الاداء التفاضلي وبين المجموعة المرجعية.

حيث يتبين من خلال النتائج السابقة المتعلقة بمعلمة الصعوبة، بأنه كلما زادت نسبة الفقرات ذات الأداء التفاضلي زادت قيمة الوسط الحسابي والانحراف المعياري لمعلمة الصعوبة. وهذه النتيجة منطقية فمن المعلوم أن الفقرات ذات الأداء التفاضلي تكون أكثر صعوبةً للمجموعة المستهدفة مقارنةً مع المجموعة المرجعية، وبالتالي كلما زادت نسبة الفقرات ذات الأداء التفاضلي في الاختبار زادت قيمة المتوسط الحسابي لمعلمة صعوبة الفقرات في الاختبار.

• فيما يخص معلمة التمييز:

تم حساب المتوسطات الحسابية والانحرافات المعيارية لمعلمة التمييز لفقرات الاختبار وفقاً لنسبة الفقرات ذات الأداء التفاضلي ولمستوى الأداء التفاضلي للفقرة، وذلك كما هو مبين في الجدول (12).

الجدول (12). الوسط الحسابي والانحراف المعياري لمعلمة التمييز لفقرات الاختبار في كل مجموعة حسب نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.

الانحراف المعياري لمعلمة التمييز	المتوسط الحسابي لمعلمة التمييز	مستوى الأداء التفاضلي للفقرة	نسبة الفقرات ذات الأداء التفاضلي
0.69	2.069	المرجعية	المرجعية
0.69	2.069	الكلية	
0.70	2.007	0.5	10%
0.49	1.757	1.5	
0.62	1.883	الكلية	
0.64	1.798	0.5	20%
0.56	1.817	1.5	
0.60	1.808	الكلية	
0.70	1.905	0.5	40%
0.55	1.591	1.5	
0.64	1.746	الكلية	
0.69	2.069	المرجعية	الكلية
0.68	1.903	0.5	
0.54	1.721	1.5	

يُلاحظ من الجدول (12)، وجود فروق ظاهرية بين المتوسطات الحسابية لمعلمة التمييز لفقرات الاختبار ناتجة عن اختلاف مجموعات متغيري الدراسة؛ حيث بلغت أعلى قيمة للوسط الحسابي لمعلمة التمييز للفقرات (2.069)، وكانت لدى المجموعة المرجعية، وأقل قيمة للوسط الحسابي لمعلمة التمييز بلغت (1.591)، وكانت لدى المجموعة المستهدفة التي تشتمل على (40% من الفقرة ذات أداء تفاضلي ومستوى الأداء التفاضلي 1.5) بفارق مقداره (0.48)، والفرق يعود لصالح المجموعة المرجعية. حيث يُلاحظ من خلال ذلك عند إشتغال الاختبار على فقرات ذات أداء تفاضلي فإن معدل التمييز لهذه الفقرات سينخفض، ويزداد مقدار الإنخفاض بزيادة مستوى الأداء التفاضلي للفقرة، وللتحقق من جوهرية الفروق الظاهرية؛ تم إجراء تحليل التباين الثنائي للمتوسطات الحسابية لمعلمة التمييز وفقاً لنسبة الفقرات ذات الأداء التفاضلي ولمستوى الأداء التفاضلي للفقرة، وذلك كما في الجدول (13).

الجدول(13): نتائج تحليل التباين الثاني للمتوسطات الحسابية لمعلمة التمييز لفقرات الاختبار حسب نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة.

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
نسبة الفقرات ذات الأداء التفاضلي	0.881	2	0.441	1.133	0.323
مستوى الأداء التفاضلي للفقرة	2.437	1	2.437	6.266	0.013
التفاعل	1.551	2	0.775	1.994	0.138
الخطأ	131.466	338	0.389		
الكل	139.174	344			

يتضح من الجدول (13)، وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين المتوسطات الحسابية لمعلمة التمييز تُعزى لمستوى الأداء التفاضلي للفقرة، أما متغير نسبة الفقرات ذات الأداء التفاضلي والتفاعل بين المتغيرين فلم تكن ذات دلالة إحصائية، ولكون المتغير مستوى الأداء التفاضلي للفقرة متعدد المجموعات؛ فقد تم إجراء اختبار Levene للتحقق من إنتهاك تجانس التباين للمتوسطات الحسابية لمعلمة التمييز لفقرات الاختبار وفقاً لمستوى الأداء التفاضلي للفقرة؛ بهدف تحديد أنسب اختبار مقارنات بعدية للكشف عن جوهرية الفروق بين المتوسطات الحسابية لمعلمة التمييز، حيث تبين عدم وجود دلالة إحصائية عند مستوى الدلالة ($\alpha=0.05$)، مما يعني عدم وجود إنتهاك لتجانس التباين لمعلمة التمييز يُعزى لمستوى الأداء التفاضلي للفقرة؛ بما يفيد ضرورة إجراء اختبار Scheffe للمقارنات البعدية المتعددة للكشف عن جوهرية الفروق بين المتوسطات الحسابية لمعلمة التمييز، وذلك كما في الجدول (14).

الجدول (14). نتائج اختبار Scheffe للكشف عن مواقع الفروق بين المتوسطات الحسابية لمعلمة التمييز لفقرات الاختبار وفقاً لمستوى الأداء التفاضلي للفقرة.

مستوى الأداء التفاضلي للفقرة	0.5	1.5	المرجعية
Scheffe	1.903	1.721	2.069
1.5			1.721
0.5		0.182	1.903
المرجعية		0.347	2.069
	0.165		

يتضح من الجدول (14)، وجود فرق دال إحصائياً عند مستوى الدلالة ($\alpha=0.05$) بين المتوسطات الحسابية لمعلمة التمييز لفقرات الاختبار وفقاً لمتغير الدراسة (مستوى الأداء التفاضلي للفقرة)؛ حيث تأثر قيمة الوسط الحسابي لمعلمة التمييز للفقرات في الاختبار المشتمل على فقرات ذات مستوى أداء تفاضلي مقداره (1.5) تأثيراً جوهرياً، بحيث انخفضت قيمة الوسط الحسابي لمعالم تمييز الفقرات في الاختبار مقارنةً مع قيمة الوسط الحسابي لمعالم تمييز الفقرات في المجموعة المرجعية، كما يُلاحظ من خلال الجدول عدم وجود فرق ذو دلالة إحصائية للأوساط الحسابية بين المجموعة المرجعية والمجموعة التي تشتمل على فقرات ذات مستوى أداء تفاضلي مقداره (0.5)، كما ويوجد فرق ذو دلالة إحصائية للأوساط الحسابية بين المجموعة التي تشتمل على فقرات ذات مستوى أداء تفاضلي مقداره (0.5) والمجموعة التي تشتمل على فقرات ذات مستوى أداء تفاضلي مقداره (1.5).

يُلاحظ من خلال بيانات الجدول (14) أنه عند إشتمال الاختبار على فقرات ذات أداء تفاضلي فإن معدل التمييز لفقرات الاختبار سينخفض، ويزداد مقدار الإنخفاض بزيادة مستوى الأداء التفاضلي للفقرة. ويعود السبب في ذلك؛ بأن الفقرات ذات الأداء التفاضلي تكون أكثر صعوبةً مما يترتب على ذلك إنخفاض معدل التمييز. حيث أشار فريسبي (Frisbie, 1973) بأن مدى الصعوبة يؤثر على الدلالة التمييزية لفقرات الاختبار، وأن الفقرات متوسطة الصعوبة يكون تمييزها أعلى منه في حالة الفقرات السهلة جداً أو الصعبة جداً، وهذا يتفق مع ما جاءت به هذه الدراسة؛ بأنه كلما زاد مستوى الأداء التفاضلي للفقرة أزداد معدل الصعوبة لفقرات الاختبار، مما ترتب على ذلك انخفاض معدل التمييز.

• فيما يخص معلمة التخمين:

تم حساب المتوسطات الحسابية والانحرافات المعيارية لمعلمة التخمين لفقرات الاختبار في كل مجموعة من مجموعات الدراسة وفقاً لمتغيري الدراسة، وذلك كما هو مبين في الجدول (15).

الجدول (15). الوسط الحسابي والانحراف المعياري لمعلمة التخمين لفقرات الاختبار في كل مجموعة حسب نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.

الانحراف المعياري لمعلمة التخمين	المتوسط الحسابي لمعلمة التخمين	مستوى الأداء التفاضلي للفقرة	نسبة الفقرات ذات الأداء التفاضلي في الاختبار
0.08	0.380	المرجعية	المرجعية
0.08	0.380	الكلية	
0.07	0.381	0.5	%10
0.06	0.342	1.5	
0.07	0.362	الكلية	
0.05	0.361	0.5	%20
0.08	0.358	1.5	
0.06	0.360	الكلية	
0.07	0.382	0.5	%40
0.07	0.332	1.5	
0.08	0.357	الكلية	
0.08	0.380	المرجعية	الكلية
0.07	0.375	0.5	
0.07	0.344	1.5	

يُلاحظ من الجدول (15)، وجود فروق ظاهرية بين المتوسطات الحسابية لمعلمة التخمين لفقرات

الاختبار ناتجة عن اختلاف مجموعات متغيري الدراسة، وللتحقق من جوهرية الفروق الظاهرية؛ تم إجراء

تحليل التباين الثنائي للمتوسطات الحسابية لمعلمة التخمين وفقاً لمتغيري الدراسة، وذلك كما في الجدول

(16).

الجدول (16). نتائج تحليل التباين الثنائي للمتوسطات الحسابية لمعلمة التخمين لفقرات الاختبار حسب نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة.

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
نسبة الفقرات ذات الأداء التفاضلي	0.001	2	0.001	0.113	0.893
مستوى الأداء التفاضلي للفقرة	0.068	1	0.068	14.415	0.000
التفاعل	0.014	2	0.007	1.513	0.222
الخطأ	1.618	338	0.005		
الكل	1.720	344			

يتضح من الجدول (16)، وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين المتوسطات الحسابية لتقديرات معلمة التخمين لفقرات الاختبار تُعزى لمستوى الأداء التفاضلي للفقرة، أما متغير نسبة الفقرات ذات الأداء التفاضلي والتفاعل بين المتغيرين فلم يكن ذو دلالة إحصائية، ولكون متغير الدراسة (مستوى الأداء التفاضلي للفقرة) متعدد المجموعات؛ فقد تم إجراء اختبار Levene للتحقق من إنتهاك تجانس التباين للمتوسطات الحسابية لمعلمة التخمين؛ بهدف تحديد أنسب اختبار مقارنات بعدية للكشف عن جوهرية الفروق بين المتوسطات الحسابية لمعلمة التخمين، حيث تبين عدم وجود دلالة إحصائية عند مستوى الدلالة ($\alpha=0.05$)، مما يعني عدم وجود انتهاك لتجانس التباين لمعلمة التخمين يُعزى لمستوى الأداء التفاضلي للفقرة؛ بما يفيد ضرورة إجراء اختبار Scheffe للمقارنات البعدية المتعددة للكشف عن جوهرية الفروق بين المتوسطات الحسابية لمعلمة التخمين لفقرات الإختبار، وذلك كما في الجدول (17).

الجدول (17). نتائج اختبار Scheffe للكشف عن مواقع الفروق بين المتوسطات الحسابية لمعلمة التخمين لفقرات الاختبار وفقاً لمستوى الأداء التفاضلي للفقرة.

مستوى الأداء التفاضلي	المتوسط الحسابي	1.5	0.5	مرجعية
Scheffe	0.344	0.344	0.375	0.380
1.5	0.344			
0.5	0.375	0.030		
مرجعية	0.380	0.036	0.006	

يتضح من الجدول (17)، وجود فرق دال إحصائياً عند مستوى الدلالة ($\alpha=0.05$) بين المتوسطات الحسابية لمعلمة التخمين وفقاً لمستوى الأداء التفاضلي للفقرة؛ حيث تأثر قيمة الوسط الحسابي لمعلمة التخمين في الاختبار المشترك على فقرات ذات مستوى أداء تفاضلي مقداره (1.5) تأثيراً جوهرياً، بحيث انخفضت قيمة الوسط الحسابي لمعالم التخمين لفقرات الاختبار مقارنة مع قيمة الوسط الحسابي لمعالم التخمين في المجموعة المرجعية، كما يُلاحظ من خلال الجدول عدم وجود فرق ذو دلالة إحصائية للأوساط الحسابية لمعلمة التخمين بين المجموعة المرجعية والمجموعة التي تشتمل على فقرات ذات مستوى أداء تفاضلي مقداره (0.5)، كما ويوجد فرق ذو دلالة إحصائية للأوساط الحسابية لمعلمة التخمين بين المجموعة التي تشتمل على فقرات ذات مستوى أداء تفاضلي مقداره (0.5)، كما ويوجد فرق ذو دلالة إحصائية للأوساط الحسابية لمعلمة التخمين بين المجموعة التي تشتمل على فقرات ذات مستوى أداء تفاضلي مقداره (1.5). حيث يُلاحظ من خلال ذلك عند إشتمال الاختبار على فقرات ذات أداء تفاضلي فإن معدل التخمين لفقرات الاختبار سينخفض ، ويزداد مقدار الإنخفاض بزيادة مستوى الأداء التفاضلي للفقرة.

وبهدف التعرف على مدى تحقق خاصية اللاتغير على مستوى المعالم الثلاث (التمييز والصعوبة والتخمين) لكل فقرة من فقرات الاختبار في كل مجموعة تبعاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة ؛ فقد تم رسم منحنى خصائص الفقرة لكل فقرة من فقرات الاختبار في كل مجموعة من المجموعات المستهدفة، بالإضافة إلى المجموعة المرجعية، والملحق (د) يوضح ذلك.

حيث تبين من خلال الشكل الموجود في الملحق (د) وجود (24) فقرة من أصل (30) فقرة من الفقرات التي لا تظهر أداء تفاضلي عبر جميع المجموعات تتطابق تقريباً منحنيات الخصائص لها؛ مما يشير إلى أن الفقرات السالفة الذكر لم تتأثر بنسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي

الفقرة، وكان هنالك (6) فقرات من أصل (30) فقرة من الفقرات التي لا تظهر أداء تفاضلي عبر جميع المجموعات لم تتطابق تقريباً منحنيات الخصائص لها، خاصةً عند الطرف الأدنى للمنحنى عند الأفراد ذوي القدرات المتدنية أما الطرف الأعلى للمنحنى عند الأفراد ذوي القدرات العالية فقد كان هنالك تطابقاً عالياً، وهذه الفقرات هي: (16,22,41,45,46,48)، حيث يُلاحظ أن منحنيات الخصائص لهذه الفقرات التي لا تظهر أداء تفاضلي عبر جميع المجموعات عند مستويات القدرة المنخفضة تأثرت بوجود الفقرات ذات الأداء التفاضلي خاصة في المجموعات ذات مستوى أداء تفاضلي مقدارة (1.5).

يعود السبب في تطابق منحنى خصائص الفقرة عند الأفراد ذوي القدرات العالية؛ بأن احتمالية الإجابة الصحيحة لهم لم تتأثر عندما إزدادت قيمة صعوبة الفقرة بسبب قدراتهم المرتفعة، كما ويعود السبب في عدم تطابق منحنى خصائص الفقرة عند الأفراد ذوي القدرات المتدنية؛ بأن هؤلاء الأفراد قد تأثرت احتمالية الإجابة الصحيحة لهم عندما إزدادت قيمة صعوبة الفقرة بسبب قدراتهم المتدنية. علماً بأن الفقرات كما هو مبين في الشكل وهي: (1,3,6,8,14,15,18,21,23,26,28,29,32,34,35,39,40,42,44,50) لم تتطابق منحنيات الخصائص لها؛ لأنها مصممة أصلاً لكي تظهر أداءً تفاضلياً.

النتائج المتعلقة بالسؤال الرابع

للإجابة عن السؤال الرابع الذي نصّ على: " هل توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين متوسطات تقديرات قدرات الأفراد على الاختبار تُعزى لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة أو التفاعل بينهما؟" تم حساب المتوسطات الحسابية والانحرافات المعيارية لتقديرات قدرات الأفراد على الاختبار وفقاً لمتغيري الدراسة، وذلك كما هو مبين في الجدول (18).

الجدول (18): الأوساط الحسابية والانحرافات المعيارية لتقديرات قدرات الأفراد على الاختبار في كل مجموعة حسب نسبة الفقرات

ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.

الانحراف المعياري للقدر	الوسط الحسابي للقدر	مستوى الأداء التفاضلي للفقرة	نسبة الفقرات ذات الأداء التفاضلي في الاختبار
1.07	0.02	صفر	المرجعية
1.07	0.02	الكلي	
1.18	-0.03	0.5	
1.12	0.01	1.5	%10
1.15	-0.01	الكلي	
1.20	-0.03	0.5	
1.14	-0.02	1.5	%20
1.17	-0.02	الكلي	
1.15	-0.02	0.5	
1.23	-0.06	1.5	%40
1.19	-0.04	الكلي	
1.07	0.02	صفر	
1.17	-0.03	0.5	الكلي
1.16	-0.02	1.5	

يُلاحظ من الجدول (18)، وجود فروق ظاهرية بين الأوساط الحسابية لتقديرات قدرات الأفراد على الاختبار ناتجة عن اختلاف مجموعات متغيري الدراسة، حيث كان أعلى وسط حسابي لقدرات الأفراد لدى المجموعة المرجعية وبلغت قيمته (0.02)، وأدنى قيمة للوسط الحسابي لقدرات الأفراد يقع لدى المجموعة المستهدفة التي تشتمل على (40% من فقرات الاختبار ذات أداء تفاضلي ومستوى الأداء التفاضلي لهذه الفقرات 1.5) وبلغت قيمته (-0.06) بفارق مقداره (0.08). ويُلاحظ من خلال الجدول أنه كلما زادت نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة قل الوسط الحسابي لقدرات الأفراد، وللتحقق من جوهرية الفروق الظاهرية؛ تم إجراء تحليل التباين الثنائي لتقديرات قدرات الأفراد على الاختبار وفقاً لمتغيري الدراسة، وذلك كما في الجدول (19).

الجدول (19). نتائج تحليل التباين الثنائي لتقديرات قدرات الأفراد على الاختبار حسب نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
نسبة الفقرات ذات الأداء التفاضلي	0.474	2	0.237	0.178	0.837
مستوى الأداء التفاضلي للفقرة	0.004	1	0.004	0.003	0.958
التفاعل	1.009	2	0.504	0.378	0.685
الخطأ	4737.704	3551	1.334		
الكلي	4740.211	3557			

يتضح من الجدول (19)، عدم وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين المتوسطات الحسابية لتقديرات قدرات الأفراد على الاختبار تعزى لنسبة الفقرات ذات الأداء التفاضلي ولمستوى الأداء التفاضلي للفقرة والتفاعل بينهما. وتتفق هذه النتيجة مع الافتراضات التي تقوم عليها نظرية الاستجابة للفقرة، وهي أن تقدير القدرة (θ) متحررة من عينة الفقرات. وكما أشار هاملتون وسواميناثان (Hambelton & Swaminathan, 1985) بأن نظرية الاستجابة للفقرة تفترض بأنه يتم الحصول على نفس التقديرات لمعالم قدرات الأفراد بغض النظر سواء كانت الفقرات سهلة أم صعبة، وهذه النتيجة تتفق مع النتيجة التي أجراها زهان وزومبو (Zhan & Zumbo, 2009). والتي تشير إلى أن الأداء التفاضلي للفقرة لا يؤثر على تقديرات قدرات الأفراد وإنما يؤثر على العلامات الملاحظة للأفراد.

كما تم حساب المتوسطات الحسابية والانحرافات المعيارية للعلامات الملاحظة للأفراد على الاختبار في كل مجموعة حسب نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة. وذلك كما هو مبين في الجدول (20).

الجدول (20). الأوساط الحسابية والانحرافات المعيارية للعلامات الملاحظة للأفراد على الاختبار في كل مجموعة وفقاً نسبة الفقرات

ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة				
الانحراف المعياري لأداء الطلبة	الوسط الحسابي لأداء الطلبة	العدد	مستوى الأداء التفاضلي للفقرة	نسبة الفقرات ذات الأداء التفاضلي
18.91	64.354	509	صفر	المرجعية
18.91	64.354	509	الكلي	
18.77	64.122	510	0.5	%10
18.05	63.917	508	1.5	
18.40	64.020	1018	الكلي	
18.58	64.126	510	0.5	%20
17.86	61.365	510	1.5	
18.27	62.745	1020	الكلي	
18.06	61.773	506	0.5	%40
16.82	57.005	505	1.5	
17.61	59.391	1011	الكلي	
18.91	64.354	509	صفر	الكلي
18.49	63.344	1526	0.5	
17.80	60.771	1523	1.5	

يُلاحظ من الجدول (20)، وجود فروق ظاهرية بين الأوساط الحسابية لأداء الأفراد على الاختبار ناتجة عن اختلاف مجموعات متغيري الدراسة، فقد كان أعلى وسط حسابي للعلامات الملاحظة للأفراد لدى المجموعة المرجعية (64.4)، وأقل وسط حسابي كان لدى المجموعة المستهدفة التي تشتمل على (40%) من فقرات الاختبار ذات أداء تفاضلي ومستوى (1.5)، وقد بلغت قيمة (57)، ويتبين من خلال ذلك أن الوسط الحسابي للعلامات الملاحظة في المجموعات يقل بشكل نمطي بزيادة نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة مقارنةً مع المجموعة المرجعية، وللتحقق من جوهرية الفروق الظاهرية؛ تم إجراء تحليل التباين الثنائي للعلامات الملاحظة للأفراد على الاختبار وفقاً لمتغيري الدراسة، وذلك كما في الجدول (21).

الجدول (21). نتائج تحليل التباين الثنائي للعلامات الملاحظة للأفراد على الاختبار وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة.

الدالة الإحصائية	قيمة ف المحسوبة	متوسط مجموع المربعات	درجة الحرية	مجموع المربعات	مصدر التباين
0.000	17.582	5800.190	2	11600.381	نسبة الفقرات ذات الأداء التفاضلي
0.000	15.355	5065.623	1	5065.623	مستوى الأداء التفاضلي للفقرة
0.018	4.024	1327.334	2	2654.668	التفاعل
		329.890	3551	1171439.996	الخطأ
			3557	1193026.849	الكل

يتضح من الجدول (21)، وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين الأوساط الحسابية للعلامات الملاحظة للأفراد على الاختبار تُعزى لنسبة الفقرات ذات الأداء التفاضلي ولمستوى الأداء التفاضلي للفقرة والتفاعل بينهما، ولكون متغيري الدراسة متعددي المجموعات؛ فقد تم إجراء اختبار Levene للتحقق من إنتهاك تجانس التباين للعلامات الملاحظة للأفراد على الاختبار؛ بهدف تحديد أنسب اختبار مقارنات بعيدية للكشف عن جوهرية الفروق بين الأوساط الحسابية للعلامات الملاحظة للأفراد على الاختبار، حيث بلغت قيمة ف المحسوبة الخاصة باختبار Levene (2.735) وهي قيمة دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) عند درجتَي حرية (6 للبسط و 3551 للمقام)، مما يعني وجود إنتهاك لتجانس التباين للعلامات الملاحظة للأفراد على الاختبار تُعزى لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة؛ بما يفيد ضرورة إجراء اختبار Games-Howell للمقارنات البعدية المتعددة للكشف عن جوهرية الفروق بين الأوساط الحسابية للعلامات الملاحظة للأفراد على الاختبار، وذلك كما في الجدولين (22) (23).

الجدول (22). نتائج اختبار Games-Howell للكشف عن مواقع الفروق للعلامات الملاحظة للأفراد على الاختبار وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار.

نسبة الفقرات ذات الأداء التفاضلي	Games-Howell	المتوسط الحسابي	%40	%20	%10	مرجعية
%40	59.391	59.391	59.391	62.745	64.020	64.354
%20	62.745	62.745	3.354	62.745	64.020	64.354
%10	64.020	64.020	4.628	1.275	64.020	64.354
مرجعية	64.354	64.354	4.962	1.608	0.334	64.354

الجدول (23). نتائج اختبار Games-Howell للكشف عن مواقع الفروق للعلامات الملاحظة للأفراد على الاختبار وفقاً لمستوى الأداء التفاضلي للفقرة.

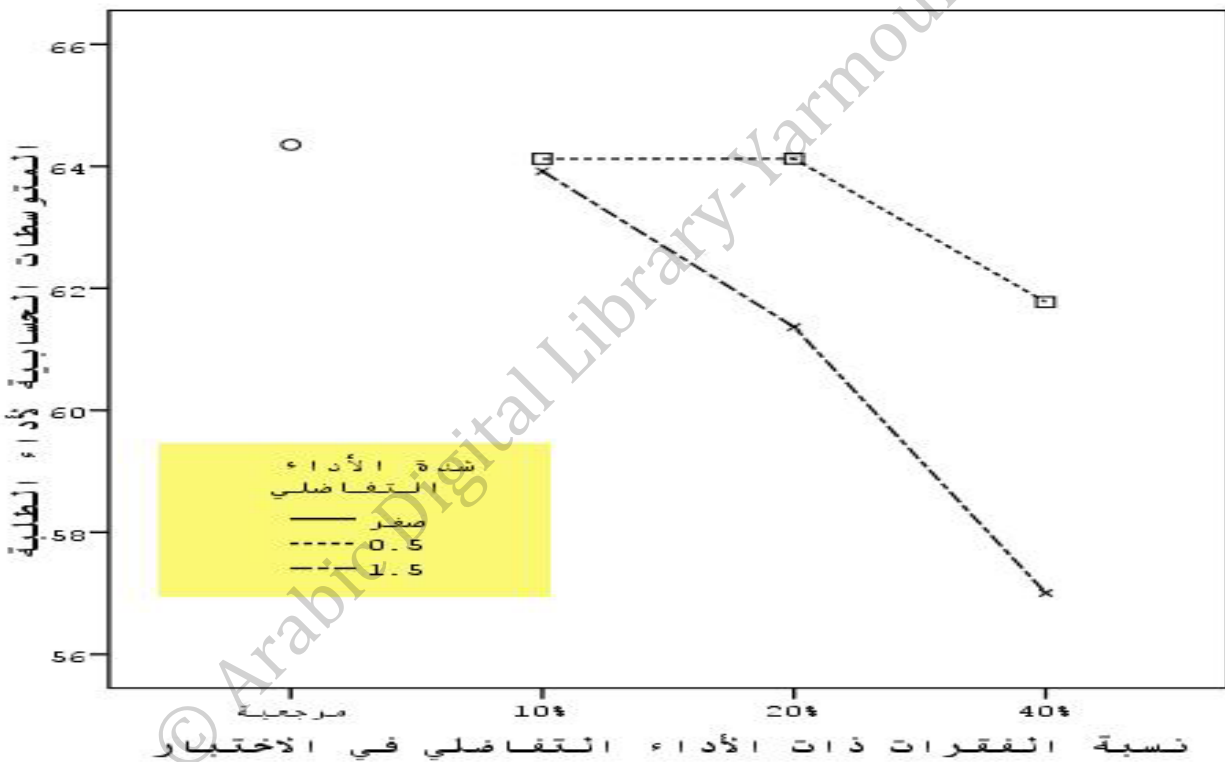
مستوى الأداء التفاضلي للفقرة	Games-Howell	المتوسط الحسابي	1.5	0.5	المرجعية
1.5	60.771	60.771	60.771	63.344	64.354
0.5	63.344	63.344	2.574	63.344	64.354
المرجعية	64.354	64.354	3.583	1.009	64.354

يتضح من الجدول (22)، وجود فرق دال إحصائياً عند مستوى الدلالة ($\alpha=0.05$) بين المتوسطات الحسابية للعلامات الملاحظة للأفراد على الاختبار وفقاً لنسبة الفقرات ذات الأداء التفاضلي، حيث تأثرت العلامات الملاحظة للأفراد على الاختبار المشتمل على 40% من الفقرات ذات الأداء التفاضلي تأثراً جوهرياً، بحيث تدنى أداؤهم مقارنةً بمن تقدموا للاختبار في كلٍّ من المجموعات: المجموعة المرجعية، والمجموعة التي تقدمت للاختبار المشتمل على 10% من الفقرات ذات الأداء التفاضلي، والمجموعة التي تقدمت للاختبار المشتمل على 20% من الفقرات ذات الأداء التفاضلي.

كما ويتضح من الجدول (23)، وجود فروق دالة إحصائياً عند مستوى الدلالة ($\alpha=0.05$) بين المتوسطات الحسابية للعلامات الملاحظة للأفراد على الاختبار وفقاً لمستوى الأداء التفاضلي للفقرة، حيث تأثرت العلامات الملاحظة للأفراد على الاختبار المشتمل على فقرات ذات مستوى أداء تفاضلي مقداره (1.5) تأثراً جوهرياً بحيث تدنت علاماتهم الملاحظة مقارنةً بمن تقدموا للاختبار في كلٍّ من المجموعات:

المجموعة المرجعية، و المجموعة التي تقدمت للاختبار المشتمل على فقرات ذات مستوى أداء تفاضلي مقداره (0.5).

وأخيراً؛ الشكل (4) يوضح التفاعل بين متغيري الدراسة (نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة) في العلامات الملاحظة للأفراد على الاختبار.



شكل (4). التفاعل بين نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة وفقاً للعلامات المشاهدة للأفراد على الاختبار.

يُلاحظ من الشكل (4)، أنه في حال الانتقال من المجموعة المرجعية إلى المجموعة المستهدفة التي تقدمت للاختبار الذي يشتمل على (10% من الفقرات ذات الأداء التفاضلي)، فالى المجموعة المستهدفة التي تقدمت للاختبار الذي يشتمل على (20% من الفقرات ذات الأداء التفاضلي) ثم المجموعة المستهدفة التي

تقدمت للاختبار الذي يشتمل على (40% من الفقرات ذات الأداء التفاضلي) يطرأ تراجع في أداء الأفراد على الاختبار على اختلاف مستوى الأداء التفاضلي للفقرة.

يتبين من خلال ذلك أن العامل الأهم في التأثير على العلامات الملاحظة للأفراد هو زيادة نسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة، فعندما كانت نسبة الفقرات ذات الأداء التفاضلي في الاختبار (20%، 10%)، ومستوى الأداء التفاضلي للفقرة (0.5) لم يكن هنالك أثر ذا دلالة احصائية على العلامات الملاحظة للأفراد، حيث ظهر الأثر عندما كانت نسبة الفقرات ذات الأداء التفاضلي في الاختبار (40%)، ومستوى الأداء التفاضلي للفقرة (1.5)، وهذه النتيجة تتفق مع دراسة زهان وزومبو (Zhan & Zumbo, 2009) التي توصلت إلى أنه كلما زاد عدد الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة تزداد الفروقات في المتوسط الحسابي للعلامة الملاحظة نتيجة لوجود الفقرات ذات الأداء التفاضلي.

النتائج المتعلقة بالسؤال الخامس

للإجابة عن السؤال الخامس الذي نص على: "هل يوجد اختلاف في كمية معلومات الاختبار يعزى لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة؟" تم حساب معاملات الثبات الأمبريقي للاختبار في كل مجموعة من المجموعات المستهدفة، بالإضافة إلى المجموعة المرجعية وفقاً لمتغيري الدراسة، ثم تم استخدام الإحصائي M للعينات المستقلة للكشف عن دلالة الفرق بين قيم معاملي الثبات الأمبريقي بين المجموعات، ويبين الجدول (24) نتائج الإحصائي M.

الجدول (24). نتائج الإحصائي M للكشف عن جوهرية الفرق بين معاملي الثبات الأمبريقي للاختبار في كل مجموعة من المجموعات المستهدفة مقارنة مع المجموعة المرجعية.

M Statistic	عدد الفقرات	العدد	الثبات الإمبريقي	حالة الفقرة في المجموعة:
0.010	50	509	0.9322	المرجعية
3.84	49	510	0.9328	المستهدفة (النسبة 10%، المستوى 0.5)
1.262	50	509	0.9322	المرجعية
3.84	50	508	0.9250	المستهدفة (النسبة 10%، المستوى 1.5)
0.105	50	509	0.9322	المرجعية
3.84	49	510	0.9302	المستهدفة (النسبة 20%، المستوى 0.5)
1.043	50	509	0.9322	المرجعية
3.84	50	510	0.9257	المستهدفة (النسبة 20%، المستوى 1.5)
0.004	50	509	0.9322	المرجعية
3.84	49	506	0.9326	المستهدفة (النسبة 40%، المستوى 0.5)
5.134	50	509	0.9322	المرجعية
3.84	48	506	0.9169	المستهدفة (النسبة 40%، المستوى 1.5)

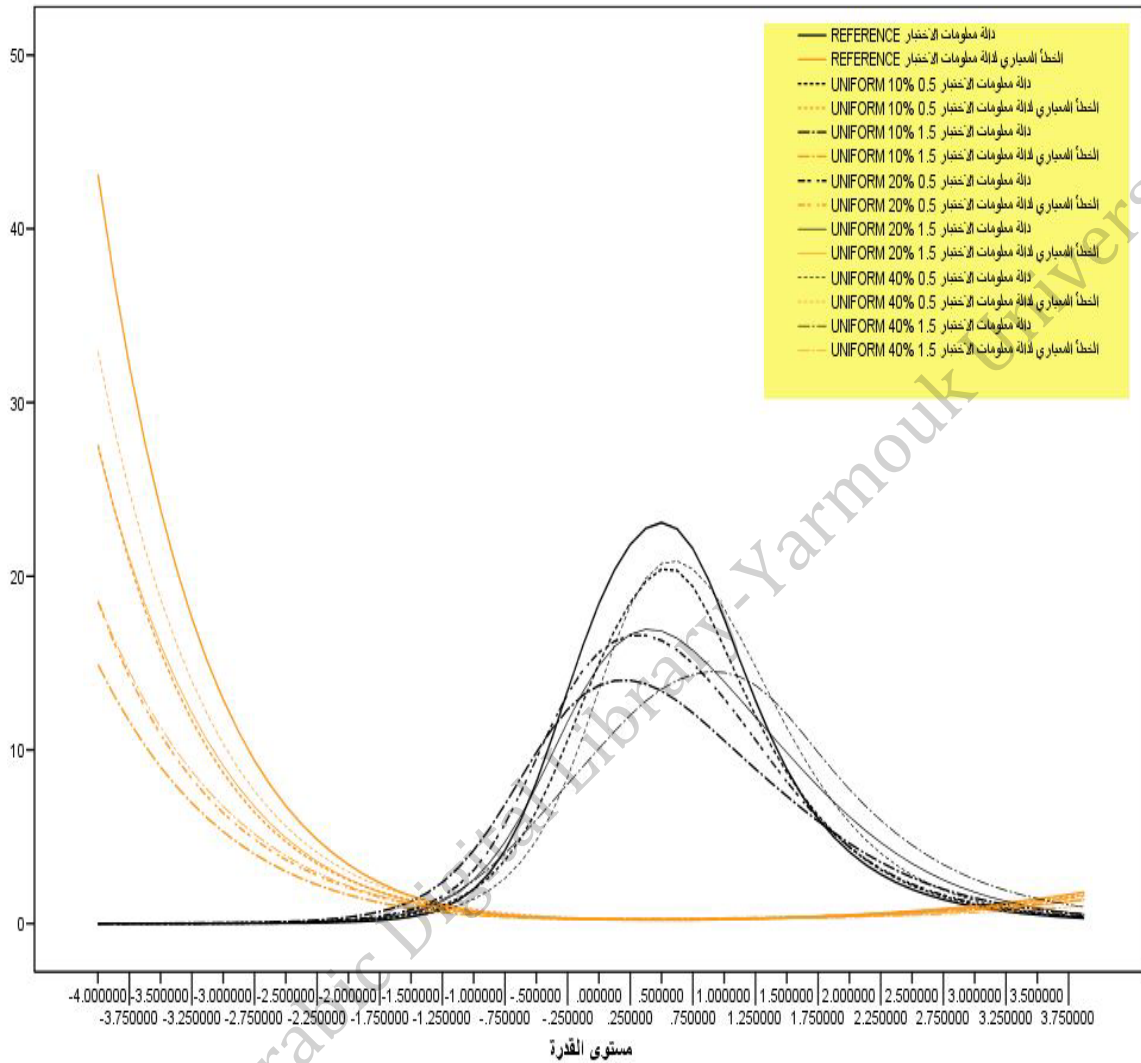
يتضح من الجدول (24)، وجود فرق دال إحصائياً عند مستوى الدلالة الإحصائية ($\alpha=0.05$) بين معاملي الثبات الأمبريقي للاختبار لدى الأفراد بين المجموعة المستهدفة التي تشتمل على (40% من الفقرات ذات أداء تفاضلي ومستوى الأداء التفاضلي 1.5) والمجموعة المرجعية يعزى لمتغيري الدراسة، حيث بلغت قيمة الإحصائي (M) المحسوبة (5.134) وهذه القيمة أكبر من قيمة (χ^2) الحرجة عند درجة حرية (5)، والبالغة قيمتها الحرجة (3.84) لصالح الثبات الأمبريقي للاختبار في المجموعة المرجعية. ويعود السبب في ذلك إلى معلمة التمييز لفقرات الاختبار، حيث أشار ايبيل وفرسبي (Ebel & Frisbie, 1986) أنه كلما زادت الدلالة التمييزية للفقرة زاد ثبات الاختبار لأن انخفاض الدلالة التمييزية للفقرة يجعل النجاح على الفقرة مرتبطاً بالصدفة، وبذلك فهي لا تقيس قدرة الطالب الحقيقية. وهذا يتفق مع جاءت به هذه الدراسة حيث تبين من خلال الإجابة عن السؤال الثالث للدراسة أن معدل التمييز لفقرات الاختبار ينخفض نتيجة لوجود الفقرات ذات الأداء التفاضلي، مما ترتب على ذلك انخفاض معامل الثبات للاختبار.

ولكون دالة معلومات الاختبار تُعدُّ مؤشراً على ثبات الاختبار كونها تتناسب عكسياً مع الخطأ المعياري للتقدير الذي يزداد ثبات الاختبار بنقصانه؛ فقد تم حساب المتوسطات الحسابية والانحرافات المعيارية لكمية معلومات الاختبار في كل مجموعة من المجموعات المستهدفة، بالإضافة الى المجموعة المرجعية وفقاً لمتغيري الدراسة، متبوعاً بالرسم البياني لمنحنى دالة معلومات الاختبار في الشكل (5).

الجدول (25). الوسط الحسابي والانحراف المعياري لكمية المعلومات التي يقدمها الاختبار في كل مجموعة وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ومستوى الأداء التفاضلي للفقرة.

حالة الفقرة في المجموعة:	المؤشر	الوسط الحسابي	الانحراف المعياري	القدرة المقابلة لأعلى دالة	الخطأ المعياري لأعلى دالة
المرجعية	دالة معلومات الاختبار	5.393	7.506	0.2081	0.500
	الخطأ المعياري لدالة معلومات الاختبار	4.992	9.421		
المستهدفة (النسبة 10%، المستوى 0.5)	دالة معلومات الاختبار	4.849	6.491	0.2214	0.500
	الخطأ المعياري لدالة معلومات الاختبار	3.447	6.047		
المستهدفة (النسبة 10%، المستوى 1.5)	دالة معلومات الاختبار	4.759	6.785	0.2672	0.250
	الخطأ المعياري لدالة معلومات الاختبار	2.135	3.333		
المستهدفة (النسبة 20%، المستوى 0.5)	دالة معلومات الاختبار	4.713	5.723	0.2455	0.375
	الخطأ المعياري لدالة معلومات الاختبار	2.555	4.102		
المستهدفة (النسبة 20%، المستوى 1.5)	دالة معلومات الاختبار	4.952	5.777	0.2429	0.375
	الخطأ المعياري لدالة معلومات الاختبار	3.452	6.143		
المستهدفة (النسبة 40%، المستوى 0.5)	دالة معلومات الاختبار	5.069	5.038	0.2189	0.625
	الخطأ المعياري لدالة معلومات الاختبار	4.023	7.259		
المستهدفة (النسبة 40%، المستوى 1.5)	دالة معلومات الاختبار	4.413	4.839	0.2627	0.875
	الخطأ المعياري لدالة معلومات الاختبار	2.565	4.229		

والشكل (5) يبين رسماً لمنحنى دالة معلومات الاختبار والخطأ المعياري في تقديرها في كل مجموعة من المجموعات المستهدفة مقارنةً مع المجموعة المرجعية وفقاً لمتغيري الدراسة.



الشكل (5). رسوم منحنيات دالة معلومات الاختبار مع الخطأ المعياري له في كل مجموعة من المجموعات المستهدفة مقارنة مع المجموعة المرجعية وفقاً لنسبة الفقرات ذات الأداء التفاضلي في الاختبار ولمستوى الأداء التفاضلي للفقرة.

يلاحظ من الجدول (25)، والشكل (5) أن النتائج الخاصة به كانت على النحو الآتي:

- فيما يخص القيمة العظمى لدالة معلومات الاختبار:

يتبين من خلال الجدول أن أعلى قيمة للوسط الحسابي والانحراف المعياري للقيمة العظمى لدالة المعلومات الاختبار كانت لدى المجموعة المرجعية، وقد بلغت القيم على الترتيب (5.39, 7.5)، وهذا يعني

أن الاختبار عند المجموعة المرجعية قدم أكبر كمية من المعلومات، مما يترتب على ذلك أن تقدير قدرات المفحوصين في المجموعة المرجعية كان الأفضل، وكانت أقل قيمة للمتوسط الحسابي والانحراف المعياري للقيمة العظمى لدالة معلومات الاختبار في المجموعة المستهدفة التي تشتمل على (40% من الفقرات ذات الأداء تفاضلي ومستوى الأداء التفاضلي 1.5)، وقد بلغت القيم على الترتيب (4.41, 4.83)، وهذا يعني أن الاختبار لدى هذه المجموعة قدم أقل كمية من المعلومات، مما يترتب على ذلك أن تقدير قدرات المفحوصين في هذه المجموعة كان الأسوأ، ويُلاحظ من خلال ذلك عند إشمال الاختبار على فقرات ذات أداء تفاضلي فإن كمية المعلومات لهذا الاختبار ستخف وتزداد مقدار الإنخفاض بازدياد مستوى الأداء التفاضلي للفقرة.

- بالنسبة للقدرة المقابلة للقيمة العظمى لدالة معلومات الاختبار، فقد قدمت المجموعة المرجعية أعلى كمية من المعلومات عند مستوى القدرة (0.20)، وهو أدنى مستوى قدرة قدم عنده أعلى كمية من المعلومات، مما يعني أن الاختبار يقيس وبشكل جيد قدرات المفحوصين الذين تقع قدراتهم عند مستوى القدرة (0.20)، وكان أعلى مستوى قدرة قدم عنده أعلى كمية من المعلومات هو مستوى القدرة (0.2627)، وذلك في المجموعة المستهدفة التي تشتمل على (40% من الفقرات ذات الأداء تفاضلي ومستوى الأداء التفاضلي 1.5).
- فيم يخص الخطأ المعياري لأعلى دالة معلومات للاختبار.

كان أعلى خطأ معياري لأعلى دالة معلومات في الاختبار عند المجموعة المستهدفة التي تشتمل على (40% من الفقرات ذات الأداء تفاضلي ومستوى الأداء التفاضلي 1.5)، حيث بلغت قيمة (0.26)، وأقل قيمة للخطأ لأعلى دالة معلومات في الاختبار كان في المجموعة المرجعية حيث بلغت قيمة (0.20)، حيث يُلاحظ من خلال ذلك عندما ازداد مستوى الأداء التفاضلي للفقرة ازدادت قيمة الخطأ المعياري لأعلى دالة معلومات للاختبار.

كما يُلاحظ من الشكل (5) أن مقدار كمية المعلومات ينخفض بازدياد مستوى الأداء التفاضلي لل فقرات، وذلك عند مستويات القدرة المتوسطة، ويرتفع مقدار كمية المعلومات بازدياد مستوى الأداء التفاضلي لل فقرات عند مستويات القدرة العليا، كما يُلاحظ وجود إزاحة الى اليمين لمنحنى دالة معلومات الاختبار في المجموعات المستهدفة مقارنةً مع المجموعة المرجعية، وذلك عند زيادة نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة، مما يترتب على ذلك أنه كلما زادت نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة ازداد مستوى القدرة الذي يقدم عنده أعلى كمية من المعلومات.

ويعود السبب في زيادة كمية معلومات الاختبار في المجموعة المرجعية؛ لأن الفقرات في هذه المجموعة أقل صعوبة مقارنةً مع باقي المجموعات المستهدفة؛ وبالتالي فإن الفقرات متوسطة الصعوبة تكون أكثر تمييزاً والفقرات الصعبة تكون أقل تمييزاً، وكما أشار (Baker,2001) بأنه كلما زاد مستوى تمييز الفقرة كلما زادت كمية المعلومات، وبناءً على ذلك كلما زادت مستوى الفقرات ذات الأداء التفاضلي كلما زادت صعوبة الفقرة مما يترتب عليه انخفاض مستوى التمييز وبالتالي انخفاض كمية معلومات الاختبار.

استنتاجات الدراسة.

لقد تناولت هذه الدراسة تقصي أثر نسبة الفقرات ذات الأداء التفاضلي في الاختبار، ومستوى الأداء التفاضلي للفقرة على تقديرات معالم فقرات الاختبار وقدرات الأفراد في ضوء النموذج اللوجستي ثلاثي المعلمة. وقد توصلت الدراسة إلى أن تقدير معالم الفقرات (الصعوبة والتمييز والتخمين)، والعلامات الملاحظة للأفراد، وكمية المعلومات التي يقدمها الاختبار، جميعها تتأثر بشكل سلبي تبعاً لنسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة، وذلك عندما كانت نسبة الفقرات ذات الأداء التفاضلي في الاختبار (40%)، ومستوى الأداء التفاضلي للفقرة (1.5)، ولم يكن هنالك أي أثر على أي من المتغيرات السابقة عندما كانت نسبة الفقرات ذات الأداء التفاضلي في الاختبار (10%, 20%)، ومستوى الأداء التفاضلي للفقرة (0.5)، لذلك لا بد لمعد الاختبار أن يأخذ بعين الاعتبار مسألة الأداء التفاضلي للفقرة، وأن تحذف الفقرات ذات الأداء التفاضلي الموجودة في الاختبار، ويعاد تقدير معالم الفقرات من جديد، وذلك عندما تزيد نسبة الفقرات ذات الأداء التفاضلي في الاختبار عن (20%)، وعندما يزيد مستوى الأداء التفاضلي للفقرة عن (0.5).

التوصيات

إن نتائج هذه الدراسة ملتزمة بنوع الفقرات ذات الأداء التفاضلي المنتظم، وبالنموذج اللوجستي ثلاثي المعلمة، ويدعو الباحث للإهتمام بزيادة الوعي حول أثر الفقرات ذات الأداء التفاضلي في معالم فقرات الاختبار وقدرات الأفراد، ويوصي بما يلي:

1. إجراء دراسة مماثلة تأخذ بعين الاعتبار الأداء التفاضلي غير المنتظم والأداء التفاضلي التفاعلي، فقد تختلف النتائج باختلاف نوع الأداء التفاضلي.
2. إجراء دراسة مماثلة تأخذ بعين الاعتبار النموذج اللوجستي إحادي وثنائي المعلمة، حيث أن الدراسة الحالية إقتصرت على النموذج اللوجستي ثلاثي المعلمة.
3. إجراء دراسة مماثلة تأخذ بعين الاعتبار حجم العينة وطول الاختبار، حيث أن الدراسة الحالية إقتصرت على حجم عينه (510) أفراد، وطول اختبار (50) فقرة في كل مجموعة.

قائمة المراجع

المراجع العربية

- إرحيل، راتب. (2011). أثر أداء الفقرات التفاضلي للجنس في دقة المعادلة العامودية لاختبار أوتيس لينون للقدرة العقلية. رسالة دكتوراه غير منشورة، جامعة اليرموك، إربد-الأردن.
- الحمادنة، إياد. (2007). الأداء التفاضلي لفقرات اختبارات تحديد الكفاءة اللغوية الإنجليزية في الجامعات الأردنية. رسالة دكتوراه غير منشورة، جامعة اليرموك، إربد-الأردن.
- الحياصات، خالد. (2011). مدى تحقق معايير الفاعلية في معادلة نموذجي اختبار مع بقاء الفقرات ذات الأداء التفاضلي للجنس وحذفها. رسالة دكتوراه غير منشورة، جامعة اليرموك، إربد-الأردن.
- عودة، أحمد. (2004). القياس والتقويم في العملية التدريسية. إربد، الأردن: دار الأمل.

المراجع الأجنبية

- Angel, M. & Gideon, d. & Jose, M. (2000). Effects of amount of dif, Test length, and purification type on Robustness and power of Mantel-Heanzel procedures. *Institute for Science Education*,5(3),44-53.
- Baker, F. (2001). *The Basics of item Response Theory*. University of Maryland, college park, MD: ERIC clearinghouse on Assessment and Evaluation.
- Berk, R. (1982). *Hand Book of Methods for Detecting Test Bias*. Baltimore, MD: The Johns Hophins University press.
- Chung, w. & Huisu, y. (2004). Effects of averge signed area between two item characteristic curves and purification procedures on the Dif detection via the mantel-hanzel method. *Applied Measurement in Education*, 17(2),113-144.
- Crocker, L. & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. New York: CBS college publishing.
- Driana, E. (2007). *Gender Diffrential item functioning on A Ninth-Grade Mathematics proficiency Test in Applachin ohio*. unpublished Doctoral Dissertation, ohio university.
- Engelhard, G. & Gamer, M. (1999). Gender difference in performance on multiple-choice and constructed response mathematics items. *Applied Measurement in Eeducation*,12(1),29-43.
- Ebel, R. & Frisbie, J. (1986). *Introduction to Classical and Modern Test Theory*. California: cole publishing company.
- Frarry, R. & Zimmer, M. (1984). Eliminating of Bias in tests Scores: Effect on Reliability and validity. *Educational and psychological Measurment*, 44(1),25-31.
- Frisbie, J. (1972). Multiple –Choice Versus True-False: A comparison of Reliability and Concurrent Validities. *Journal of Educational Measurement*,10(4),247-304.

- Gillian, p. (1989). Examining the Relationship between Differential Item functioning and Item Difficulty. *College Board Report No. 89-5*.
- Hambleton, R. & Rogers, J. (1995). Item bias review practical assessment, research and evaluation. Retrieved July 31, 2003 from [http://edresearch.org/pare/getvn.Asp=4 & n=6](http://edresearch.org/pare/getvn.Asp=4&n=6).
- Hambleton, R. & Rogers, J. (1989). Detecting potentially bias test item: comparison of IRT aria and mantel-haenszel methods. *Applied Measurement in Education*,2(4),313-334.
- Hambleton, R. & Swaminathan, H. (1985). *Item Response theory: principles and applications*. Boston, MA: Kluwek NIJ publishing.
- Hambleton, R. & Swaminathan, H. & Rogers, J. (1991). *Fundamentals of item response theory*. Newbury perk California: stage publications.
- HO Kim, S. & Cohen, A. (1995). Acomparision of Lord chi-square, Raju's area measures, and the Likelihood ratio test on detection of differential item functioning. *Applied Measurement in Education*,8(4),291-312.
- Jensen, R. (1980). *Bias in mental testing*. new York: Macmillan publishing co. Inc.
- John, n. & William, D. & Barbara, M. (2002). Impact of Differential Item functioning on Age and Gender Differences in functional Disability. *Applied Measurement in Education*,15(3),120-154.
- Jodoin, m. & Gierl , M. & Ackerman, T. (2000). Performance of mantel-Haenszel , simultane ous item bias test , and logisitic regression when the proportion of Dif item is large. *paper presented at the annual meeting of the American educational Researh Association Orleans, Louisiana, April,24-27,2000*.
- Kamata, A. & Chu, k. (2003). Test equating in the presence of Dif items. *Journal of Applied Measurement*,6(3), 342.
- Kamp, L. & Gruijter, D. (2005). Statistical test theory for education and psychology. Retrived December 30, 2005 from: [www. Leidenuniv. nl/ gruibterdnmde](http://www.Leidenuniv.nl/gruiterdnmde).

- Linn, L. & Harnisch, D. (1981). interactions between item content and group membership in achievement test item. *Journal of Educational measurement*, 18(2),109-118.
- Lord, F. (1980). *Application of item Response Theory to practical Testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lord, F. & Novick, M. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-wesley.
- Luppescu, S. (2002). DIF detection in HLM. *Paper presented at the annual meeting of the American Educational Research Association*, April, 2000, New Orleans, LA.
- Osterlind, S. (1983). *Test item bias*. Beverly Hills: sage publications.
- pae, T. (2004). Dif for examinees with different academic backgrounds. *Language Testing*,21(1),53-72.
- Popham, J. (1999). *Classroom Assessment: What teachers need to know*. Boston: Allyn and Bacon.
- Raju, N. (1988). The area between two item characteristic curves. *Psychometrika*, 53,495-502.
- Raju, S. & Ellis , B. (2003). Test and item bias: what they are, what arent, and How to detect them. *Measuring up. Assessment issues for teachers, counselors and administrator*. (ERIC Document Reproduction service NO,ED480042).
- Risner, A. (2008). *Effect of unequal sample size on the power of Dif Detection*. unpublished Doctoral Dissertation, the faculty of the Virginia polytechnic institute and state university.
- Samejima, F. (1977). A use Of the information function in tailored testing. *Applied psychological Measurement*, 1,233-247.
- Swaminathan, H. & Rogers, J. (1990). Detecting differential item functioning using Logistic regression procedures. *Journal of Educational Measurement* ,127(4).

- Subkoviak, M. & Grag, R. & Mack, J. & Ironson, G. (1984). Empirical comparison of selected item Bias Detection procedures with Bias manipulation. *Journal of Educational measurement*, 21(1), 49-58.
- Tan, x. & Jeral, M. (2005). using Local Dif analyses to assess group differences on multilingual examinations. *Applied Measurement in Education*, 2, 313-334.
- Teresi, J. (2001). Statistical methods of examination of differential item functioning with applications to cross-clutural measurement of functional, physical and mental health. *J mental Health and Aging*, 7, 31-40.
- Thurman, C. (2009). *A monte carlo Study investigating the influence of the Discrimination , Category intersection parameters , and differential item functioning patterns on the Detection of Dif in polytomous item*. un published Doctoral Dissertation, Georgia state university.
- Walsted, W. & Robson, D. (1997). Item functioning and male-female differences on multiple-choice test in Economics. *The Journal of Economic Education*, 28(2), 155-178.
- Zhang, y. & Matthews, J. & Dorans, N. (2003). *Using dif dissection to assess effects of item deletion due to dif on the performance of sat I: Reasoning test subpopulations (Report)*. u.s, New Jersey.
- Zhen, L. & Zumbo, B. (2009). Impact of Differential Item Functioning on Subsequent statistical conclusions Based on observed Test Score Data. *Psicologica*, 30(2), 343-370.
- Zumbo, B. (2003). Does item-Level Dif manifest itself in scale-Level analyses? Implications for translating language tests. *Language Testing*, 20(2), 136-147.
- Zumbo, B. (2000). The Effect of DIF and Impact on Classical Test Statistics: Undetected DIF and Impact, and the Reliability and Interpretability of Scores from a Language Proficiency Test. *Presented at the annual conference of the nationl counic on measurement in Education (NCME)*, April, 2000 new Orleans, LA.

الملحق (أ)

قيم معالم الفقرة (الصعوبة، التمييز، التخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المرجعية.

رقم الفقرة	معلمة التمييز	الخطأ المعياري لمعلمة التمييز	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة	معلمة التخمين	الخطأ المعياري لمعلمة التخمين	χ^2	درجة الحرية	احتمالية الخطأ
1	0.971	0.273	0.732	0.346	0.436	0.071	4.10	8	0.845
2	1.043	0.314	1.076	0.318	0.500	0.053	7.50	8	0.481
3	1.522	0.442	0.921	0.206	0.486	0.048	7.90	8	0.440
4	1.068	0.292	1.004	0.257	0.337	0.061	12.30	8	0.139
5	1.078	0.290	0.446	0.338	0.478	0.071	5.40	8	0.710
6	1.322	0.338	0.827	0.206	0.380	0.054	3.40	8	0.907
7	1.268	0.402	1.269	0.272	0.500	0.044	3.50	8	0.896
8	2.015	0.463	0.727	0.133	0.400	0.042	0.60	8	1.000
9	1.725	0.450	0.989	0.152	0.375	0.043	0.40	8	1.000
10	1.407	0.312	0.504	0.194	0.347	0.058	7.80	8	0.450
11	2.499	0.658	0.917	0.115	0.449	0.036	5.50	8	0.708
12	1.877	0.422	0.388	0.175	0.482	0.050	2.90	7	0.890
13	1.452	0.348	0.728	0.180	0.355	0.052	7.80	8	0.455
14	1.392	0.324	0.683	0.185	0.331	0.055	10.20	8	0.250
15	2.005	0.443	0.549	0.146	0.444	0.045	13.50	8	0.097
16	1.284	0.328	0.552	0.248	0.438	0.061	5.70	8	0.685
17	1.871	0.393	0.138	0.189	0.488	0.054	9.50	7	0.216
18	1.141	0.303	0.609	0.279	0.500	0.056	3.60	8	0.888
19	1.806	0.385	0.508	0.153	0.386	0.049	11.80	8	0.159
20	2.006	0.438	0.540	0.144	0.421	0.046	2.30	8	0.970
21	2.632	0.578	0.541	0.121	0.478	0.040	8.20	8	0.418
22	1.298	0.286	-0.203	0.318	0.468	0.079	8.70	7	0.274
23	1.768	0.378	0.290	0.182	0.434	0.055	8.40	7	0.296
24	1.407	0.312	0.855	0.155	0.243	0.048	10.20	8	0.249
25	1.415	0.287	0.162	0.215	0.365	0.063	12.70	8	0.121
26	2.944	0.707	0.727	0.106	0.470	0.036	6.00	7	0.541
27	1.584	0.290	-0.312	0.226	0.405	0.070	3.00	7	0.881
28	2.023	0.349	-0.083	0.156	0.379	0.056	11.90	7	0.105
29	2.745	0.512	0.116	0.125	0.478	0.045	11.70	7	0.111
30	2.284	0.358	0.064	0.117	0.303	0.047	7.60	7	0.368
31	1.864	0.318	0.196	0.138	0.294	0.051	10.40	8	0.240
32	2.832	0.512	0.289	0.108	0.401	0.042	6.70	7	0.460
33	2.440	0.389	-0.042	0.119	0.346	0.048	2.30	7	0.940
34	2.947	0.617	0.728	0.089	0.357	0.035	4.40	7	0.728
35	2.212	0.372	0.187	0.123	0.336	0.047	1.70	7	0.975
36	1.937	0.315	0.065	0.137	0.305	0.051	7.90	7	0.345
37	2.610	0.441	-0.043	0.122	0.398	0.049	4.50	7	0.726
38	2.477	0.432	0.064	0.127	0.400	0.048	7.90	7	0.345
39	2.007	0.343	0.365	0.119	0.284	0.045	3.80	8	0.873
40	2.578	0.386	0.339	0.087	0.229	0.037	8.80	7	0.271
41	2.692	0.487	0.407	0.104	0.361	0.041	4.90	7	0.672
42	2.170	0.356	0.413	0.104	0.255	0.041	5.60	8	0.693
43	3.616	0.637	0.696	0.067	0.263	0.030	6.80	7	0.452
44	2.471	0.375	-0.466	0.128	0.304	0.055	17.00	7	0.018
45	3.388	0.525	-0.124	0.087	0.320	0.043	12.40	6	0.054
46	2.602	0.422	-0.129	0.119	0.369	0.049	6.30	7	0.510
47	3.053	0.474	-0.058	0.095	0.337	0.044	6.00	7	0.538
48	3.013	0.474	-0.287	0.103	0.327	0.048	6.40	6	0.380
49	3.352	0.498	0.450	0.068	0.221	0.032	3.10	7	0.875
50	2.312	0.375	-0.202	0.135	0.360	0.053	6.60	7	0.473

المحلق (ب)

قيم معالم الفقرة (الصعوبة، التمييز، التخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المستهدفة التي تشتمل على (10% من الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة 0.5).

رقم الفقرة	معلمة التمييز	الخطأ المعياري لمعلمة التمييز	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة	معلمة التخمين	الخطأ المعياري لمعلمة التخمين	χ^2	درجة الحرية	احتمالية الخطأ
1	0.725	0.238	1.698	0.581	0.500	0.060	8.3	8	0.405
2	0.868	0.283	1.206	0.422	0.500	0.062	2.1	8	0.978
3	1.188	0.341	0.530	0.319	0.500	0.066	6.3	8	0.614
4	1.303	0.400	1.042	0.221	0.424	0.051	11.2	8	0.192
5	1.898	0.503	0.613	0.175	0.500	0.045	7.2	8	0.515
6	1.749	0.588	1.158	0.170	0.461	0.040	5.3	8	0.727
7	0.986	0.258	0.434	0.359	0.404	0.079	8.0	8	0.431
8	1.212	0.394	1.176	0.247	0.425	0.052	8.5	8	0.389
9	1.346	0.338	0.815	0.183	0.308	0.053	5.1	8	0.750
10	1.183	0.362	1.025	0.244	0.403	0.057	13.6	8	0.094
11	1.729	0.437	0.675	0.166	0.409	0.048	16.2	8	0.040
12	1.055	0.251	0.070	0.374	0.432	0.084	9.2	8	0.328
13	1.620	0.414	0.789	0.160	0.366	0.047	7.0	8	0.539
14	1.384	0.450	1.134	0.214	0.445	0.047	6.5	8	0.589
15	1.678	0.421	0.643	0.175	0.415	0.049	6.6	8	0.578
16	1.917	0.488	0.728	0.147	0.419	0.044	8.4	8	0.392
17	1.565	0.356	0.146	0.247	0.465	0.064	1.8	8	0.986
18	1.605	0.342	0.069	0.236	0.449	0.064	9.2	8	0.323
19	1.595	0.348	0.453	0.180	0.357	0.054	1.6	8	0.990
20	1.303	0.266	0.178	0.241	0.339	0.068	16.2	8	0.040
21	1.316	0.244	0.013	0.237	0.309	0.071	23.6	8	0.003
22									
23	1.329	0.294	0.077	0.283	0.417	0.073	6.9	8	0.543
24	2.151	0.554	0.943	0.110	0.340	0.037	8.8	8	0.356
25	1.203	0.227	-0.095	0.275	0.336	0.077	11.3	8	0.188
26	1.799	0.421	0.584	0.160	0.395	0.048	1.9	8	0.983
27	1.577	0.270	-0.560	0.254	0.403	0.082	12.7	8	0.122
28	1.879	0.358	-0.058	0.200	0.428	0.061	12.8	8	0.120
29	2.081	0.378	-0.021	0.171	0.397	0.056	21.6	8	0.006
30	2.395	0.454	0.329	0.129	0.385	0.043	6.0	8	0.642
31	1.977	0.405	0.510	0.136	0.336	0.045	7.6	8	0.475
32	2.624	0.532	0.405	0.124	0.428	0.041	8.8	8	0.358
33	1.783	0.370	0.179	0.196	0.427	0.056	4.4	8	0.824
34	1.795	0.325	0.491	0.127	0.237	0.044	14.9	8	0.061
35	2.769	0.444	0.340	0.095	0.263	0.037	8.9	8	0.354
36	1.753	0.327	0.119	0.182	0.354	0.058	9.2	8	0.325
37	1.829	0.288	-0.366	0.188	0.342	0.067	12.8	8	0.120
38	1.619	0.282	-0.253	0.222	0.368	0.071	11.3	8	0.184
39	4.126	0.700	0.586	0.063	0.280	0.029	6.3	8	0.614
40	2.659	0.495	0.432	0.109	0.355	0.039	3.1	8	0.925
41	2.356	0.448	0.576	0.105	0.293	0.038	11.5	8	0.176
42	2.738	0.478	0.599	0.086	0.264	0.033	2.9	8	0.941
43	2.271	0.465	0.813	0.093	0.247	0.034	8.2	8	0.413
44	2.464	0.400	-0.266	0.145	0.377	0.055	6.1	8	0.636
45	2.502	0.418	0.018	0.131	0.378	0.046	7.5	8	0.489
46	3.013	0.541	-0.147	0.130	0.446	0.048	14.7	7	0.041
47	2.753	0.440	-0.011	0.116	0.338	0.045	14.8	8	0.063
48	2.512	0.408	-0.294	0.144	0.385	0.055	13.5	8	0.094
49	2.557	0.395	0.372	0.093	0.219	0.036	5.0	8	0.762
50	3.623	0.655	-0.035	0.104	0.420	0.041	3.8	7	0.801

(الملحق ت)

قيم معالم الفقرة (الصعوبة، التمييز، التخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المستهدفة التي تشتمل على (10% من الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة 1.5).

رقم الفقرة	معلمة التمييز	الخطأ المعياري لمعلمة التمييز	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة	معلمة التخمين	الخطأ المعياري لمعلمة التخمين	χ^2	درجة الحرية	احتمالية الخطأ
1	1.512	0.487	1.115	0.204	0.457	0.046	18.9	9	0.026
2	0.715	0.212	0.895	0.523	0.424	0.084	9.5	9	0.393
3	1.012	0.267	0.658	0.329	0.402	0.073	10.2	9	0.337
4	1.187	0.329	0.895	0.256	0.390	0.061	7.7	9	0.568
5	0.857	0.190	0.029	0.405	0.363	0.089	7.4	9	0.599
6	1.179	0.361	1.086	0.265	0.425	0.057	9.0	9	0.441
7	1.096	0.348	1.141	0.301	0.449	0.059	11.0	9	0.276
8	1.009	0.269	0.696	0.325	0.388	0.073	4.3	9	0.887
9	1.065	0.362	1.475	0.325	0.432	0.053	3.0	9	0.965
10	1.187	0.316	0.809	0.258	0.386	0.063	4.3	9	0.887
11	1.000	0.229	0.348	0.325	0.356	0.078	9.0	9	0.442
12	0.942	0.207	-0.144	0.404	0.393	0.092	5.7	9	0.774
13	1.052	0.256	0.823	0.258	0.298	0.065	6.3	9	0.707
14	1.494	0.412	0.886	0.202	0.417	0.052	11.1	9	0.270
15	1.151	0.244	0.291	0.273	0.351	0.072	8.3	9	0.504
16	1.224	0.205	0.033	0.217	0.264	0.067	6.0	9	0.736
17	1.264	0.226	-0.248	0.262	0.352	0.078	15.3	9	0.083
18	1.198	0.253	0.217	0.275	0.374	0.073	5.1	9	0.826
19	1.351	0.229	0.137	0.198	0.265	0.064	11.1	9	0.269
20	1.915	0.394	0.470	0.167	0.394	0.051	9.8	9	0.363
21	1.342	0.241	0.108	0.218	0.319	0.067	4.1	9	0.908
22	1.751	0.285	-0.257	0.183	0.366	0.067	3.3	9	0.949
23	1.612	0.253	-0.295	0.187	0.312	0.069	10.6	9	0.307
24	1.250	0.295	0.869	0.201	0.280	0.056	8.6	9	0.476
25	1.588	0.264	-0.094	0.194	0.338	0.067	10.7	9	0.298
26	1.465	0.254	0.018	0.206	0.327	0.067	3.0	9	0.964
27	2.251	0.364	-0.448	0.147	0.385	0.064	4.9	8	0.764
28	1.685	0.278	-0.451	0.203	0.356	0.076	19.5	9	0.021
29	1.917	0.331	0.297	0.151	0.312	0.053	8.0	9	0.538
30	1.447	0.246	-0.041	0.208	0.315	0.068	4.2	9	0.895
31	1.626	0.294	0.271	0.179	0.326	0.058	4.8	9	0.855
32	1.401	0.221	-0.210	0.204	0.277	0.070	11.1	8	0.197
33	1.775	0.290	-0.121	0.178	0.337	0.065	11.1	9	0.271
34	1.316	0.255	0.420	0.206	0.296	0.061	8.1	9	0.523
35	1.782	0.322	0.668	0.129	0.234	0.043	7.1	9	0.631
36	1.686	0.302	0.150	0.185	0.354	0.061	3.4	9	0.947
37	1.564	0.242	-0.542	0.199	0.312	0.075	10.2	9	0.333
38	2.096	0.314	-0.330	0.141	0.327	0.061	2.8	9	0.973
39	1.962	0.621	1.340	0.131	0.273	0.034	9.4	9	0.398
40	2.287	0.333	0.252	0.111	0.255	0.044	5.2	9	0.819
41	1.709	0.249	0.054	0.145	0.244	0.055	9.2	9	0.422
42	1.899	0.607	1.328	0.137	0.295	0.035	6.8	9	0.662
43	1.701	0.345	0.685	0.149	0.286	0.047	7.3	9	0.606
44	2.403	0.385	-0.554	0.134	0.361	0.064	3.3	8	0.913
45	2.183	0.336	-0.208	0.139	0.347	0.059	3.9	9	0.920
46	2.304	0.404	-0.229	0.159	0.461	0.060	3.2	8	0.920
47	2.365	0.332	-0.205	0.113	0.280	0.053	10.1	9	0.343
48	2.743	0.442	-0.399	0.115	0.373	0.057	6.5	8	0.587
49	2.631	0.426	0.448	0.105	0.277	0.039	16.5	9	0.057
50	2.383	0.356	-0.468	0.122	0.306	0.060	4.4	8	0.821

المحلق (ث)

قيم معالم الفقرة (الصعوبة، التمييز، التخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المستهدفة التي تشتمل على (20% من الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة 0.5).

رقم الفقرة	معلمة التمييز	الخطأ المعياري لمعلمة التمييز	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة	معلمة التخمين	الخطأ المعياري لمعلمة التخمين	χ^2	درجة الحرية	احتمالية الخطأ
1									
2	1.071	0.320	0.887	0.319	0.471	0.063	9.2	9	0.422
3	0.851	0.254	0.975	0.427	0.500	0.065	1.4	9	0.998
4	1.101	0.325	1.148	0.264	0.367	0.058	5.9	8	0.662
5	1.035	0.238	0.282	0.322	0.379	0.077	5.8	8	0.665
6	1.373	0.400	1.037	0.215	0.426	0.050	10.8	9	0.287
7	1.084	0.304	0.876	0.288	0.410	0.064	4.7	9	0.859
8	1.086	0.334	1.280	0.283	0.391	0.056	2.5	9	0.980
9	1.068	0.257	0.620	0.271	0.322	0.069	19.0	8	0.015
10	1.257	0.333	0.791	0.239	0.409	0.058	3.1	8	0.929
11	1.433	0.361	0.767	0.198	0.381	0.054	5.8	9	0.762
12	1.006	0.207	-0.246	0.368	0.377	0.090	8.4	8	0.398
13	2.310	0.609	1.037	0.118	0.399	0.035	3.4	8	0.910
14	1.238	0.310	0.656	0.242	0.377	0.062	5.6	9	0.781
15	1.889	0.453	0.705	0.154	0.419	0.045	7.4	8	0.491
16	1.760	0.386	0.474	0.174	0.415	0.051	7.8	9	0.550
17	1.114	0.238	-0.236	0.362	0.415	0.089	7.0	8	0.541
18	1.256	0.282	0.225	0.273	0.403	0.071	5.9	9	0.746
19	1.028	0.216	0.097	0.316	0.343	0.080	7.5	8	0.484
20	1.316	0.300	0.452	0.232	0.366	0.064	15.7	9	0.073
21	1.549	0.323	0.414	0.186	0.350	0.056	8.5	9	0.487
22	1.266	0.245	-0.508	0.326	0.420	0.090	1.4	9	0.998
23	2.292	0.409	0.075	0.142	0.399	0.050	5.2	8	0.738
24	1.752	0.374	0.719	0.141	0.312	0.045	6.6	8	0.581
25	1.413	0.257	-0.079	0.227	0.344	0.070	7.7	8	0.465
26	1.141	0.234	0.200	0.267	0.331	0.072	4.3	9	0.890
27	1.706	0.298	-0.342	0.211	0.386	0.072	9.7	9	0.378
28	1.837	0.289	-0.389	0.176	0.335	0.067	7.0	8	0.532
29	1.968	0.359	0.241	0.150	0.362	0.050	5.0	8	0.759
30	1.719	0.299	0.016	0.177	0.332	0.061	14.7	9	0.100
31	2.140	0.432	0.456	0.137	0.388	0.045	8.5	8	0.383
32	2.015	0.331	0.026	0.146	0.318	0.054	16.3	8	0.038
33	1.874	0.285	-0.148	0.150	0.289	0.058	6.9	8	0.546
34	1.926	0.429	0.737	0.133	0.338	0.044	6.2	9	0.722
35	2.535	0.542	0.702	0.108	0.371	0.037	10.9	8	0.210
36	2.106	0.321	-0.047	0.132	0.282	0.053	18.6	8	0.017
37	2.660	0.404	-0.332	0.118	0.338	0.053	7.9	8	0.444
38	2.405	0.394	-0.220	0.140	0.371	0.056	4.5	8	0.812
39	1.600	0.323	0.689	0.144	0.258	0.047	3.3	8	0.912
40	1.990	0.316	0.243	0.124	0.257	0.046	4.0	8	0.855
41	1.981	0.377	0.421	0.138	0.337	0.047	3.2	9	0.956
42	2.101	0.469	0.945	0.108	0.274	0.036	5.9	8	0.662
43	2.579	0.561	0.870	0.096	0.329	0.034	5.5	8	0.703
44	2.233	0.314	-0.307	0.124	0.266	0.055	7.7	8	0.464
45	2.548	0.438	-0.110	0.136	0.410	0.052	8.4	8	0.392
46	2.767	0.462	-0.338	0.129	0.404	0.056	5.8	8	0.672
47	1.894	0.291	-0.225	0.157	0.299	0.061	10.0	8	0.262
48	3.103	0.497	-0.423	0.108	0.363	0.053	8.8	6	0.186
49	3.422	0.570	0.348	0.081	0.298	0.035	3.1	8	0.926
50	3.307	0.545	-0.193	0.102	0.365	0.047	4.5	8	0.808

المعلق (ج)

قيم معالم الفقرة (الصعوبة، التمييز، التخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المستهدفة التي تشتمل على (20% من الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة 1.5).

رقم الفقرة	معلمة التمييز	الخطأ المعياري لمعلة التمييز	معلمة الصعوبة	الخطأ المعياري لمعلة الصعوبة	معلمة التخمين	الخطأ المعياري لمعلة التخمين	χ^2	درجة الحرية	احتمالية الخطأ
1	1.341	0.466	1.519	0.281	0.439	0.043	5.4	9	0.794
2	0.998	0.326	1.277	0.363	0.500	0.055	2.8	9	0.972
3	0.957	0.242	0.351	0.367	0.395	0.082	7.4	9	0.597
4	1.585	0.437	0.880	0.180	0.429	0.048	3.7	8	0.883
5	0.803	0.239	0.525	0.518	0.493	0.084	14.5	9	0.107
6	1.578	0.544	1.515	0.236	0.395	0.038	5.0	8	0.755
7	1.305	0.381	0.796	0.252	0.500	0.054	6.3	9	0.713
8	1.801	0.571	1.169	0.169	0.453	0.039	2.4	9	0.984
9	1.266	0.277	0.500	0.211	0.300	0.062	10.0	8	0.265
10	0.998	0.324	1.317	0.343	0.446	0.058	5.8	9	0.764
11	1.811	0.529	0.972	0.163	0.464	0.042	3.0	9	0.965
12	1.153	0.289	0.295	0.318	0.455	0.072	3.3	9	0.950
13	1.520	0.333	0.595	0.169	0.312	0.053	2.7	8	0.954
14	1.979	0.597	1.031	0.152	0.485	0.038	3.9	9	0.917
15	1.272	0.308	0.469	0.247	0.393	0.065	1.4	8	0.995
16	1.915	0.425	0.636	0.141	0.375	0.045	6.7	8	0.573
17	1.420	0.308	0.068	0.256	0.419	0.070	9.0	8	0.345
18	1.202	0.298	0.627	0.239	0.350	0.064	5.1	8	0.743
19	1.319	0.305	0.435	0.229	0.365	0.063	3.7	8	0.883
20	2.165	0.469	0.599	0.132	0.415	0.042	4.7	8	0.790
21	2.089	0.463	0.803	0.115	0.321	0.039	3.3	8	0.911
22	1.294	0.247	-0.202	0.271	0.361	0.078	13.7	9	0.132
23	1.383	0.267	0.103	0.219	0.326	0.066	16.6	8	0.035
24	2.110	0.470	0.848	0.111	0.304	0.038	3.7	8	0.882
25	1.795	0.417	0.473	0.180	0.468	0.050	5.6	9	0.780
26	1.375	0.267	0.274	0.197	0.291	0.061	8.0	8	0.437
27	1.656	0.261	-0.392	0.192	0.318	0.068	7.8	8	0.451
28	1.728	0.328	-0.153	0.212	0.414	0.066	6.3	8	0.618
29	1.610	0.310	0.698	0.128	0.208	0.044	9.7	8	0.286
30	1.971	0.361	0.221	0.148	0.356	0.051	1.8	8	0.986
31	2.130	0.370	0.294	0.126	0.320	0.045	4.4	8	0.820
32	1.939	0.383	0.213	0.166	0.421	0.052	3.8	9	0.927
33	2.177	0.346	-0.025	0.132	0.310	0.051	10.4	8	0.241
34	1.642	0.339	0.489	0.161	0.323	0.052	8.4	8	0.391
35	2.149	0.606	1.309	0.124	0.253	0.031	6.1	8	0.641
36	2.026	0.361	0.137	0.148	0.349	0.052	10.7	8	0.220
37	1.967	0.322	-0.307	0.170	0.347	0.063	13.7	8	0.090
38	2.033	0.320	-0.364	0.161	0.330	0.062	7.6	8	0.477
39	1.879	0.641	1.725	0.238	0.225	0.029	5.9	8	0.655
40	2.279	0.386	0.329	0.114	0.312	0.043	0.8	8	0.999
41	2.272	0.419	0.453	0.117	0.334	0.042	5.8	8	0.665
42	1.527	0.548	1.937	0.334	0.243	0.032	5.1	8	0.750
43	1.943	0.431	0.826	0.119	0.300	0.040	3.1	8	0.929
44	2.309	0.349	0.179	0.106	0.257	0.042	2.0	8	0.981
45	3.137	0.489	0.028	0.092	0.316	0.041	2.3	7	0.942
46	2.405	0.368	-0.355	0.132	0.322	0.055	4.6	8	0.795
47	2.512	0.383	-0.190	0.118	0.301	0.050	9.6	8	0.297
48	2.662	0.405	-0.501	0.123	0.313	0.056	12.2	7	0.094
49	3.153	0.483	0.395	0.078	0.250	0.034	3.6	8	0.894
50	3.329	0.548	-0.179	0.097	0.341	0.045	10.3	7	0.171

المعلق (ح)

قيم معالم الفقرة (الصعوبة، التمييز، التخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المستهدفة التي تشتمل على (40% من الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة 0.5).

رقم الفقرة	معلمة التمييز	الخطأ المعياري لمعلمة التمييز	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة	معلمة التخمين	الخطأ المعياري لمعلمة التخمين	χ^2	درجة الحرية	احتمالية الخطأ
1	1.452	0.493	1.315	0.238	0.486	0.042	2.2	8	0.973
2	1.184	0.381	1.133	0.286	0.500	0.052	4.6	8	0.803
3	0.897	0.282	0.976	0.412	0.495	0.068	10.7	8	0.217
4	2.545	0.888	1.193	0.128	0.497	0.033	4.8	8	0.780
5	1.545	0.353	0.208	0.231	0.445	0.061	11.0	8	0.201
6	1.607	0.503	1.197	0.183	0.423	0.042	3.8	8	0.875
7	0.798	0.250	1.075	0.476	0.483	0.072	11.3	8	0.185
8	1.509	0.552	1.599	0.281	0.438	0.038	8.1	8	0.419
9	1.203	0.322	0.873	0.222	0.344	0.058	7.0	8	0.533
10	1.712	0.468	0.950	0.161	0.429	0.043	13.4	8	0.100
11	1.673	0.436	0.837	0.166	0.424	0.045	1.8	8	0.987
12	1.543	0.405	0.521	0.223	0.500	0.054	3.7	8	0.887
13	1.821	0.478	0.901	0.150	0.423	0.042	1.5	8	0.992
14	1.693	0.466	0.969	0.161	0.413	0.043	2.5	8	0.962
15	1.472	0.311	0.534	0.170	0.286	0.053	14.7	8	0.065
16	1.693	0.454	0.874	0.165	0.432	0.045	3.4	8	0.908
17	1.128	0.245	-0.080	0.338	0.399	0.083	9.4	8	0.312
18	1.205	0.367	1.052	0.249	0.430	0.054	2.1	8	0.978
19	1.566	0.348	0.576	0.168	0.344	0.051	3.0	8	0.933
20	1.494	0.291	0.080	0.211	0.343	0.063	7.6	8	0.478
21	2.048	0.370	0.567	0.109	0.250	0.040	15.4	8	0.051
22	1.316	0.254	-0.495	0.327	0.414	0.090	5.4	8	0.709
23	1.969	0.404	0.472	0.142	0.388	0.045	3.8	8	0.875
24	1.919	0.408	0.823	0.116	0.283	0.039	6.1	8	0.637
25	2.229	0.393	0.156	0.129	0.347	0.046	6.0	8	0.650
26	2.049	0.518	0.893	0.133	0.422	0.039	4.9	8	0.772
27	1.596	0.313	-0.190	0.241	0.401	0.071	13.5	8	0.097
28	2.302	0.438	0.080	0.144	0.422	0.048	4.6	8	0.797
29	1.973	0.402	0.491	0.139	0.384	0.044	7.2	8	0.517
30	2.519	0.452	0.338	0.112	0.386	0.041	8.5	8	0.389
31	1.638	0.324	0.352	0.165	0.324	0.052	11.2	8	0.193
32	1.737	0.307	0.070	0.166	0.303	0.055	17.6	8	0.024
33	2.061	0.362	-0.068	0.155	0.343	0.053	7.1	8	0.523
34	2.642	0.489	0.755	0.088	0.290	0.034	3.7	8	0.885
35	2.011	0.419	0.796	0.112	0.289	0.039	12.8	8	0.119
36	2.819	0.506	0.349	0.105	0.417	0.039	7.6	8	0.469
37									
38	2.186	0.442	-0.061	0.176	0.466	0.053	6.8	8	0.559
39	1.507	0.336	0.910	0.138	0.227	0.043	11.2	8	0.192
40	2.912	0.489	0.603	0.084	0.298	0.034	6.1	8	0.641
41	3.245	0.597	0.622	0.087	0.385	0.034	11.3	8	0.187
42	2.531	0.733	1.203	0.102	0.318	0.032	10.8	8	0.213
43	3.191	0.617	0.871	0.076	0.287	0.031	2.6	8	0.957
44	2.923	0.501	0.131	0.100	0.372	0.040	9.0	8	0.343
45	2.046	0.361	-0.167	0.168	0.353	0.057	8.8	8	0.358
46	3.180	0.656	-0.030	0.120	0.492	0.042	8.4	6	0.208
47	3.394	0.613	0.012	0.093	0.376	0.039	15.1	7	0.035
48	1.808	0.311	-0.456	0.216	0.369	0.073	19.4	8	0.013
49	2.779	0.394	0.380	0.076	0.204	0.032	2.7	8	0.951
50	4.071	0.656	0.289	0.071	0.365	0.035	7.8	7	0.354

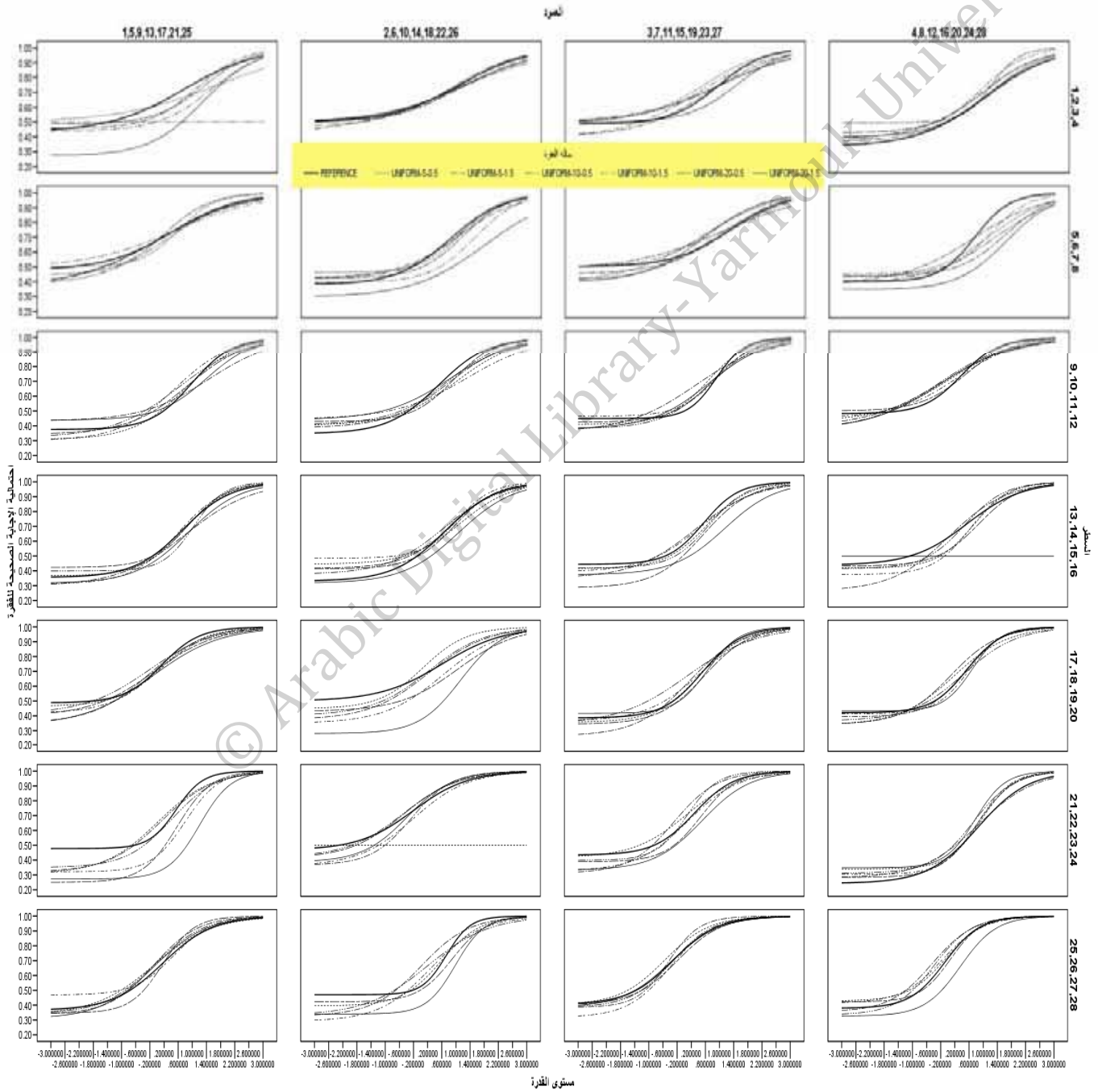
(الملحق خ)

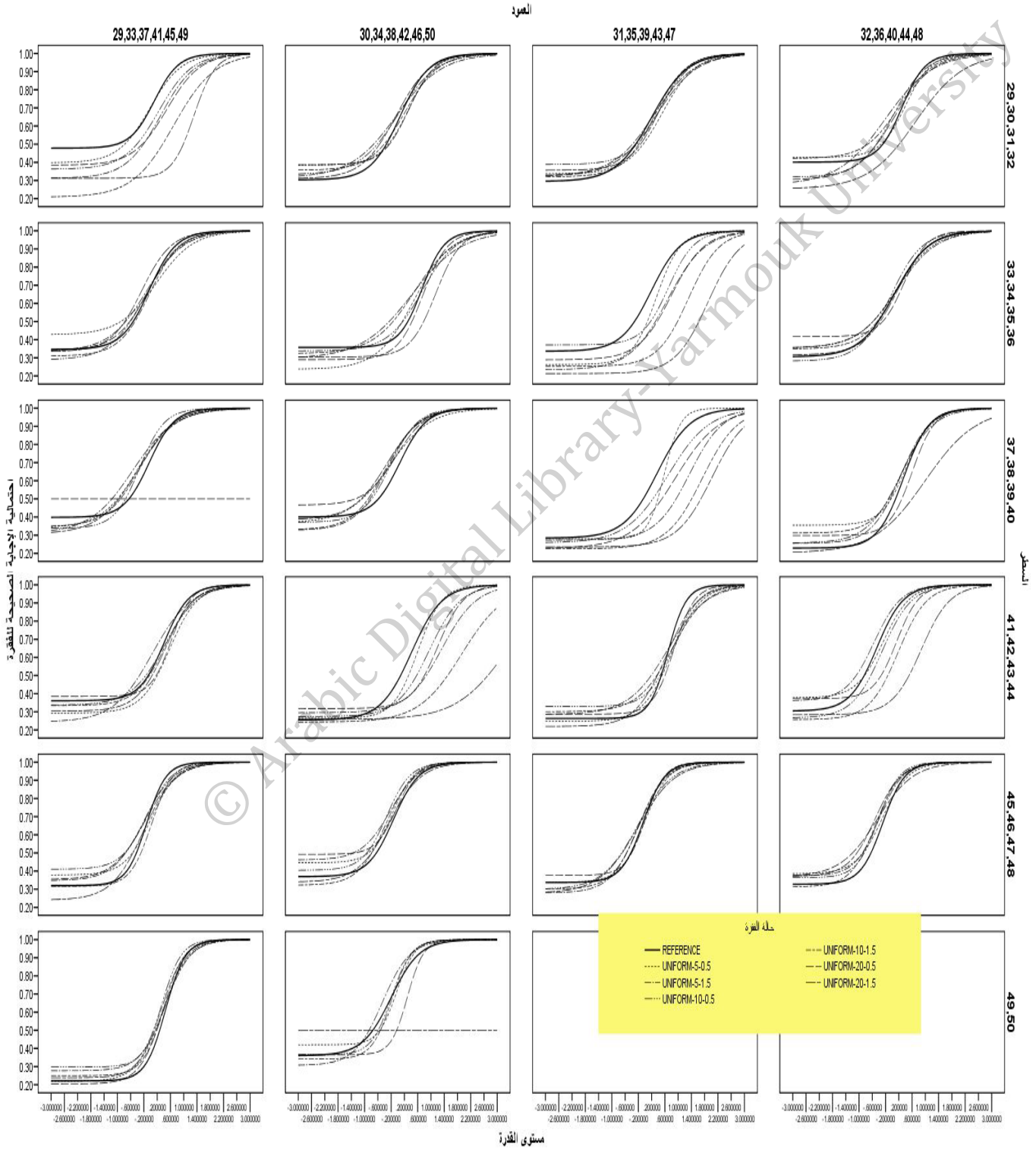
قيم معالم الفقرة (الصعوبة، التمييز، التخمين) والخطأ المعياري، والاختبار الإحصائي χ^2 لحسن المطابقة في المجموعة المستهدفة التي تشتمل على (40% من الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة 1.5).

رقم الفقرة	معلمة التمييز	الخطأ المعياري لمعلمة التمييز	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة	معلمة التخمين	الخطأ المعياري لمعلمة التخمين	χ^2	درجة الحرية	احتمالية الخطأ
1	1.320	0.382	1.287	0.188	0.272	0.045	4.0	8	0.854
2	0.873	0.285	1.223	0.425	0.500	0.062	2.1	8	0.978
3	1.464	0.535	1.449	0.264	0.500	0.039	4.6	9	0.867
4	1.269	0.398	1.159	0.226	0.398	0.051	4.9	9	0.843
5	0.877	0.206	0.057	0.430	0.376	0.093	12.0	8	0.150
6	0.983	0.328	1.813	0.357	0.298	0.050	5.8	8	0.675
7	1.192	0.308	0.540	0.273	0.397	0.068	5.8	9	0.763
8	1.459	0.513	1.613	0.257	0.346	0.039	1.9	8	0.984
9	1.474	0.511	1.402	0.232	0.439	0.041	11.2	8	0.192
10	1.100	0.331	0.842	0.298	0.440	0.065	5.0	8	0.753
11	1.228	0.331	0.709	0.244	0.379	0.063	7.4	8	0.489
12	0.990	0.210	-0.183	0.385	0.375	0.092	5.5	9	0.792
13	1.346	0.351	0.883	0.186	0.316	0.053	11.6	9	0.235
14	1.276	0.358	1.080	0.197	0.317	0.051	5.9	8	0.663
15	1.376	0.423	1.153	0.198	0.374	0.048	3.3	9	0.953
16									
17	1.038	0.207	-0.130	0.333	0.336	0.086	4.9	8	0.766
18	1.696	0.462	1.131	0.136	0.281	0.039	10.0	8	0.262
19	2.234	0.526	0.661	0.130	0.417	0.040	2.3	8	0.970
20	2.463	0.590	0.696	0.120	0.432	0.038	9.5	9	0.391
21	2.236	0.610	1.205	0.109	0.273	0.031	8.3	8	0.403
22	1.573	0.279	-0.367	0.244	0.387	0.078	3.0	8	0.936
23	1.466	0.333	0.509	0.194	0.333	0.058	9.9	9	0.362
24	2.364	0.539	0.775	0.107	0.348	0.037	5.5	9	0.793
25	1.328	0.238	-0.109	0.242	0.309	0.074	9.7	8	0.285
26	2.453	0.631	0.986	0.098	0.340	0.034	5.6	9	0.782
27	1.445	0.263	-0.378	0.270	0.385	0.082	15.0	8	0.059
28	1.969	0.380	0.361	0.145	0.326	0.048	5.9	8	0.656
29	3.081	0.981	1.305	0.101	0.312	0.028	5.0	8	0.755
30	1.964	0.321	-0.056	0.160	0.310	0.058	12.7	8	0.124
31	2.286	0.447	0.411	0.128	0.356	0.043	4.4	9	0.887
32	1.392	0.304	0.716	0.165	0.253	0.051	11.6	8	0.169
33	2.193	0.341	-0.276	0.152	0.334	0.059	11.8	9	0.227
34	2.624	0.713	1.140	0.092	0.305	0.030	7.9	8	0.443
35	2.024	0.769	1.898	0.306	0.213	0.026	10.8	8	0.215
36	2.059	0.354	0.123	0.147	0.316	0.052	4.2	8	0.836
37	1.714	0.263	-0.401	0.192	0.308	0.071	16.3	8	0.038
38	1.916	0.323	-0.334	0.194	0.386	0.068	6.3	8	0.618
39	1.912	0.768	2.017	0.372	0.235	0.026	2.2	8	0.975
40	1.298	0.302	0.997	0.160	0.202	0.047	10.9	8	0.205
41	2.307	0.433	0.452	0.117	0.304	0.041	3.2	8	0.922
42	1.320	0.561	3.264	1.168	0.258	0.026	8.8	8	0.361
43	2.143	0.394	0.640	0.101	0.219	0.037	4.8	8	0.780
44	2.297	0.514	0.874	0.097	0.284	0.034	6.0	8	0.647
45	2.075	0.285	-0.216	0.132	0.240	0.053	10.2	8	0.251
46	2.085	0.330	-0.249	0.160	0.339	0.060	18.3	9	0.032
47	1.769	0.267	-0.225	0.169	0.279	0.063	6.1	8	0.637
48	2.510	0.406	-0.521	0.149	0.372	0.065	11.6	7	0.115
49	2.870	0.445	0.330	0.089	0.238	0.035	3.1	8	0.925
50									

الملحق (د)

رسوم منحني خصائص الفقرة لكل فقرة من فقرات الاختبار في كل مجموعة من المجموعات المستهدفة بالإضافة إلى المجموعة المرجعية وفقاً لمتغيري الدراسة (نسبة الفقرات ذات الأداء التفاضلي ومستوى الأداء التفاضلي للفقرة).





Abstract

Al Nawafleh, Ali Saleh. The Effect of percentage of items with Differential Functioning and its Level on Estimating Item parameters and persons According to the Three parameters model. PhD dissertation, Yarmouk University, 2013 (Supervisor Professor: Ahmad Qawasmeh).

The study aimed to identify the effect of percentage and level of Items with Differential Function on the Test Items Characteristics and Individual's Abilities based on 3PL using a stimulated data study. The study used different test conditions based on the two conditions used in the current study (percentage and level items with Differential Function). percentage of Items with Differential Function used in the current study is (10%,20%,40%). the level of Items with Differential Function were (low, high) representing (0.5,1.5). The researcher also used a reference group that does not contain any item has Differential Function, accordingly , resulting in (7) study data sets.

The researcher used (Wingen) statistical software to stimulate the responses (510) examinees in each group, and the total data generated was (3570) examinees. Then, (Bilog MG) was used to verify the unidimensionality of the resulting and the fit of items and examinees, items and examinees parameters estimation, calculating the test information function and the standard error in ability estimation according to each of the individual data sets. Results of the study indicated:

1. Significant differences were found at ($\alpha=0.05$) in item difficulty estimation due to the variables of the study (percentage of Items with Differential Function).
2. Significant differences were found at ($\alpha=0.05$) in item discrimination and guessing estimation due to the variable of the study (level of items with Differential Function) in favor of the reference group.

3. There were no significant differences at ($\alpha=0.05$) in examinees abilities estimation due to the variables of the study (percentage and level of Items with Differential Function), and Significant differences were found at ($\alpha=0.05$) in the observed scores of examinees due to the variables of the study (percentage and level of Items with Differential Function) and the interaction in favor of the reference group.
4. There were significant differences in the means scores of item information function due to the percentage and level of Items with Differential Function in favor of the reference group.

Key words: percentage of Items with Differential Function, level of Items with Differential Function, reference group, focal group, Items Parameters, Test Information Function, Ability.